Model Identification and Data Analysis

Matteo Secco

June 24, 2021

Contents

Part I
Prediction

1 Probability Recall

1.1 Random Vectors

Variance $Var[v] = E[(v - E[v])^2]$

Cross-Variance $Var[v, u] = E[(v - E[v])(u - E[u])]$

Variance Matrix $\begin{vmatrix} Var[v_1] & . & . & Var[v_1, v_k] \\ . & . & & . \\ . & & . & . \\ Var[v_k, v_1] & . & . & Var[v_k] \end{vmatrix}$

Covariance coefficient $\delta[i, j] = \frac{Var[i,j]}{\sqrt{Var[i]}\sqrt{Var[j]}}$
$\delta[i, j] = 0 \implies$ i, j uncorrelated
$|\delta[i, j]| = 1 \implies i = \alpha j$

1.2 Random processes

$v(t, s)|$ t time instant, s expetiment outcome (generally given)

Mean $m(t) = E[v(t, s)]$

Variance $\lambda^2(t) = Var[v(t)]$

Covariance function $\gamma(t_1, t_2) = E[(v(t_1) - m(t_1))(v(t_2) - m(t_2))] = \gamma(t_2, t_1)$

Normalized Covariance Function $\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)}$
$\forall$ stationary processes: $|\rho(\tau)| \leq 1 \quad \forall\tau$

1.3 Important process classes

Stationary process

- $m(t) = m$ constant
- $\lambda^2(t) = \lambda^2$ constant
- $\gamma(t_1, t_2) = f(t_2 - t_1) = \gamma(\tau)$ covariance depends only on time difference τ

$|\gamma(\tau)| \leq \gamma(0) \quad \forall\tau$

White noise $\eta(t) \sim WN(m, \lambda^2)$

- Stationary process
- $\gamma(\tau) = 0 \quad \forall \tau \neq 0$

$v(t) = \alpha\eta(t) + \beta \quad \eta(t) \sim WN(0, \lambda^2) \qquad \Longrightarrow \qquad v(t) \sim WN(\beta, \alpha^2\lambda^2)$

2 Spectral Analysis

2.1 Foundamentals

Spectrum

$$\Gamma(\omega) = \overbrace{F(\gamma(\tau))}^{\text{Fourier transform}} = \sum_{\tau=-\infty}^{+\infty} \gamma(\tau) \cdot e^{-j\omega\tau}$$

Euler formula $\Gamma(\omega) = \gamma(0) + 2cos(\omega)\gamma(1) + 2cos(2\omega)\gamma(2) + ...$

Spectrum properties

- $\Gamma : \mathbb{R} \to \mathbb{R}$
- Γ is periodic with $T = 2\pi$
- Γ is even $[\Gamma(-\omega) = \Gamma(\omega)]$
- $\Gamma(\omega) \geq 0 \quad \forall\omega$

$\eta(t) \sim WN(0, \lambda^2) \quad \implies \quad \Gamma_\eta(\omega) = \gamma(0) = Var[\eta(t)] = \lambda^2$

Anti-Transform

$$\gamma(\tau) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \Gamma(\omega) e^{k\omega\tau} \, dw$$

Complex spectrum

$$\phi(z) = \sum_{\tau=-\infty}^{+\infty} \omega(\tau) z^{-\tau}$$

$$\Gamma(\omega) = \Phi(e^{j\omega})$$

2.2 Fundamental theorem of Spectral Analysis

Fundamental theorem of Spectral Analysis allows to derive the (real and/or complex) spectrum of the output from the input and the transfer function of the system

$$\xrightarrow{u} \boxed{W(z)} \xrightarrow{y}$$

$$\Gamma_{yy}(\omega) = |W(e^{j\omega})|^2 \cdot \Gamma_{uu}(\omega)$$
$$\Phi_{yy}(z) = W(z)W(z^{-1}) \cdot \Phi_{uu}(z)$$

2.3 Canonical representation of a Stationary Process

A stationary process can be represented by an infinite number of transfer functions. The canonical representation is the transfer function $W(z)$ such that:

- Numerator and denominator have same degree
- Numerator and denominator are monic (highest grade coefficient is 1)
- Numerator and denominator are coprime ($W(z)$ cannot be simplified)
- numerator and denominator are stable polynomials (all poles and zeros of $W(z)$ are inside the unit disk)

3 Moving Average Processes

Given $\eta(t) \sim WN(0, \lambda^2)$

3.1 MA(1):

Model

$$v(t) = c_0\eta(t) + c_1\eta(t-1)$$

Mean

$$\begin{aligned}
E[v(t)] &= c_0 \cdot E[\eta(t)] + c_1 \cdot E[\eta(t)] \\
&= c_0 \cdot 0 \qquad + c_1 \cdot 0
\end{aligned}$$

$$\boxed{E[v(t)] = 0}$$

Variance

$$\begin{aligned}
Var[v(t)] &= E[(v(t) \underbrace{-E[v(t)]}_{0})^2] \\
&= E[(v(t))^2] \\
&= E[(c_0 \cdot \eta(t)^2 \qquad + c_1 \cdot \eta(t-1))^2] \\
&= c_0^2 \cdot E[\eta(t)^2] \qquad + c_1^2 \cdot E[\eta(t-1)^2] + \underbrace{2c_0c_1 \cdot E[\eta(t)\eta(t-1)]}_{0} \\
&= c_0^2 \cdot E[\eta(t)^2] \qquad + c_1^2 \cdot E[\eta(t-1)^2] \\
&= c_0^2\lambda^2 \qquad + c_1^2\lambda^2
\end{aligned}$$

$$\boxed{Var[v(t)] = (c_0^2 + c_1^2)\lambda^2}$$

Covariance

$$\begin{aligned}
\gamma(t_1, t_2) &= E[(v(t_1) - E[v(t_1)]) \qquad \cdot (v(t_2) - E[v(t_2)])] \\
&= E[(c_0\eta(t_1) + c_1\eta(t_1 - 1)) \ \cdot (c_0\eta(t_2) + c_1\eta(t_2 - 1))] \\
&= c_0^2 E[\eta(t_1)\eta(t_2)] \qquad + c_1^2 E[\eta(t_1 - 1)\eta(t_2 - 1) \\
&\qquad + c_0c_1E[\eta(t_1)\eta(t_2 - 1)] + c_0c_1E[\eta(t_1 - 1)\eta(t_2)]
\end{aligned}$$

$$\boxed{\gamma(\tau) = \begin{cases} c_0^2\lambda^2 + c_1^2\lambda^2 & \text{if } \tau = 0 \\ c_0c_1\lambda^2 & \text{if } \tau = \pm 1 \\ 0 & \text{otherwise} \end{cases}}$$

3.2 MA(n)

Model

$$\begin{aligned} v(t) &= c_0\eta(t) + c_1\eta(t-1) + ... + c_n\eta(t-n) \\ &= (c_0 + c_1 z^{-1} + ... + c_n z^{-n})\eta(t) \end{aligned}$$

Transfer function

$$W(z) = c_0 + c_1 z^{-1} + ... + c_n z^{-n} = \frac{c_0 z^n + c_1 z^{n-1} + ... + c_n}{z^n}$$

All poles are in the complex origin

Mean

$$E[v(t)] = (c_0 + c_1 + ... + c_n)\underbrace{E[\eta(t)]}_{0}$$

$$\boxed{E[v(t) = 0}$$

Covariance function

$$\boxed{\gamma(\tau) = \begin{cases} \lambda^2 \cdot \sum_{i=0}^{n-\tau} c_i c_{i-\tau} & |\tau| \leq n \\ 0 & \text{otherwise} \end{cases}}$$

example

$$\begin{aligned} \gamma(0) &= (c_0^2 + c_1^2 + ... + c_n^2)\lambda^2 \\ \gamma(1) &= (c_0 c_1 + c_1 c_2 + ... + c_{n-1} c_n)\lambda^2 \\ \gamma(2) &= (c_0 c_2 + c_1 c_3 + ... + c_{n-2} c_n)\lambda^2 \\ &... \\ \lambda(n) &= (c_0 c_n)\lambda^2 \\ \lambda(k) &= 0 \,\forall k > n \end{aligned}$$

3.3 MA(∞)

Model

$$v(t) = c_0\eta(t) + c_1\eta(t-1) + ... + c_k\eta(t-k) + ... = \sum_{i=0}^{\infty} c_i\eta(t-i)$$

Variance

$$\gamma(0) = (c_0^2 + c_1^2 + ... + c_k^2 + ...)\lambda^2 = \lambda^2 \sum_{i=0}^{\infty} c_i^2$$

3.4 Well definition of an MA(∞)

We need to have $|\gamma(\tau)| \leq \gamma(0)$, so we must require that

$$\boxed{\gamma(0) = \lambda^2 \sum_{i=0}^{\infty} c_i^2 \text{ is finite}}$$

4 Auto Regressive Processes

4.1 AR(1)

Model

$$v(t) = av(t-1) + \eta(t)$$

Mean

$$\begin{aligned}
E[v(t)] &= E[av(t-1)] + \overbrace{E[\eta(t)]}^{0} \\
&= aE[v(t-1)] \\
&= aE[v(t)] \\
(1-a)E[v(t)] &= 0 \\
\boxed{E[v(t)] = 0}
\end{aligned}$$

Covariance

MA(∞) method Observe as an AR(1) can be axpressed as an MA(∞)

$$\begin{aligned}
v(t) &= av(t-1) && +\eta(t) \\
&= a[av(t-2) + \eta(t-1)] && +\eta(t) \\
&= a^2 v(t-2) && +a\eta(t-1) + \eta(t) \\
&= a^2[v(t-3) + \eta(t-2)] && +a\eta(t-1) + \eta(t) \\
&= \underbrace{a^n v(t-n)}_{\to 0} + \underbrace{\sum_{i=0}^{\infty} a^i \eta(t-i)}_{\text{MA}(\infty)}
\end{aligned}$$

In particular, the result depends on an $MA(\infty)$ having $\sum_{i=0}^{\infty} c_i = \sum_{i=0}^{\infty} a^i$. To check if the variance is finite we check $\gamma(0) = \lambda^2 \sum_{i=0}^{\infty} a^{2i} < \infty$. The given is a geometric series, convergent for $|a| < 1$. Under this hypothesis its value is

$$\gamma(0) = \lambda^2 \sum_{i=0}^{\infty} a^{2i} = \frac{\lambda^2}{1-a^2}$$

Applying the formula of the variance of MA processes we get

$$\gamma(1) = (c_0c_1 + c_1c_2 + ...)\lambda^2 = (a + aa^2 + ...)\lambda^2 = a(1 + a^2 + a^4 + ...)\lambda^2 = a\lambda^2 \sum_{i=0}^{\infty} a^{2i} = a\frac{\lambda^2}{1-a^2} = a\gamma(0)$$

$$\gamma(2) = (c_0c_2 + c_1c_3 + ...)\lambda^2 = (a^2 + aa^3 + ...)\lambda^2 = a^2(1 + a^2 + a^4 + ...)\lambda^2 = a^2\lambda^2 \sum_{i=0}^{\infty} a^{2i} = a^2\frac{\lambda^2}{1-a^2} = a^2\gamma(0)$$

$$\boxed{\gamma(\tau) = a^{|\tau|}\frac{\lambda^2}{1-a^2}}$$

Yule-Walkler Equations

$$\begin{aligned}
Var[v(t)] &= E[v(t)^2] \\
&= E[(av(t) + \eta(t))^2] \\
&= a^2 \underbrace{E[v(t-1)^2]}_{\substack{=Var[v(t-1)] \\ =Var[v(t)] \\ =\gamma(0)}} + \underbrace{E[\eta(t)^2]}_{=\lambda^2} + 2a \underbrace{E[v(t-1)\eta(t)]}_{\substack{v(t-1) \text{ depends on } \eta(t-2) \\ \eta(t) \text{ independent of } \eta(t-2) \\ \Longrightarrow \\ E[v(t-1)\eta(t)]=0}}
\end{aligned}$$

$$\gamma(0) = a^2\gamma(0) + \lambda^2$$

$$\boxed{\gamma(0) = \frac{\lambda^2}{1-a^2}}$$

To find $\gamma(\tau)$, we start from the model $v(t) = av(t-1) + \eta(t)$.

$$\begin{aligned}
v(t) &= av(t-1) & &+\eta(t) \\
v(t)v(t-\tau) &= av(t-1)v(t-\tau) & &+\eta(t)v(t-\tau) \\
\underbrace{E[v(t)v(t-\tau)]}_{\gamma(\tau)} &= a\,\underbrace{E[v(t-1)v(t-\tau)]}_{\gamma(\tau-1)} & &+ \underbrace{E[\eta(t)v(t-\tau)]}_{0}
\end{aligned}$$

$$\boxed{\gamma(\tau) = a\gamma(\tau-1)}$$

We can join the two by inductive reasoning, obtaining

$$\boxed{\gamma(\tau) = a^{|\tau|}\frac{\lambda^2}{1-a^2}}$$

Long Division Leads to same result, but is boring

4.2 AR(n)

Model

$$v(t) = a_1v(t-1) + a_2v(t-2) + ... + a_nv(t-n) + \eta(t)$$

Transfer function

$$W(z) = \frac{z^n}{z^n - a_1 z_{n-1} - ... - a_n}$$

Mean

$$E[v(t)] = a_1 E[v(t-1]) + a_2 E[v(t-2)] + ... + a_n E[v(t-n)] + \underbrace{E[\eta(t)]}_{0}$$

$$m = a_1 m \qquad + a_2 m \qquad + ... + a_n m$$

$$(1 - a_1 - a_2 - ... - a_n) m = 0$$

$$\boxed{E[v(t)] = 0}$$

5 ARMA Processes

Model

$$v(t) = a_1 v(t-1) + ... + a_{n_a} v(t-n_a) + c_0 \eta(t) + ... + c_{n_c} v(t-n_c)$$

Can also be espressed as $V(t) = \frac{C(z)}{A(z)} \eta(t)$, where

$$C(z) = c_0 + c_1 z^{-1} + ... + c_{n_c} z^{-n_c}$$
$$A(z) = 1 - a_1 z^{-1} - ... - a_{n_a} z^{-n_a}$$

Such process is stationary if all the poles of $W(z)$ are inside the unit disk.

6 Prediction problem

We want to predict $v(t+r)$ from $v(t), v(t-1), ...$, where r is called prediction horizon, of the following stationary process:

$$\xrightarrow{\ \eta\ } \boxed{W(z)} \xrightarrow{\ v\ }$$

6.1 Fake problem

Having a process with transfer function $W(z)$, we can compute it in polynomial form using the long division algorithm

$$W(z) = w_0 + w_1 z^{-1} + w_2 z^{-2} + ...$$

We can calculate

$$v(t+r) = W(z)\eta(t+r) = \underbrace{w_0\eta(t+r) + w_1\eta(t+r-1) + ... + w_{r-1}\eta(t+1)}_{\alpha(t)\ \textbf{unpredictable:}\ \text{future of } \eta \text{ involved}} + \underbrace{w_r\eta(t) + w_{r+1}\eta(t-1) + ...}_{\beta(t)\ \textbf{predictable}}$$

The optimal fake predictor is then

$$\boxed{v(t+r|t) = w_r\eta(t) + w_{r+1}\eta(t-1) + ...} = \beta(t)$$

And the prediction error is

$$\begin{aligned} \epsilon(t) &= v(t+r) && -\hat{v}(t+r|t) \\ &= \alpha(t) + \beta(t) && -\beta(t) \\ &= \alpha(t) \end{aligned}$$

$$\boxed{\epsilon(t) = w_0\eta(t+r) + w_1\eta(t+r-1) + ... + w_{r-1}\eta(t+1)}$$

$$\boxed{Var[\epsilon(t)] = (w_0^2 + w_1^2 + ... + w_{r-1}^2)\lambda^2}$$

6.2 True Problem

We want to estimate $v(t+r)$ form $v(t)$, having transfer function $W(z)$ and $\hat{W}_r(z)$ the solution to the fake problem. We can calculate the transfer function of the real predictor from the process as

$$\boxed{W_r(z) = W(z)^{-1} \cdot \hat{W}_r(z)}$$

For ARMA processes a shortcut exists:

$$\hat{v}_{\mathrm{ARMA}}(t|t-1) = \frac{C(z)A(z)}{C(z)} \qquad \text{having } W(z) = \frac{C(z)}{A(z)}$$

6.3 Prediction with eXogenous variables

An exogenous variable is a deterministic input variable in the system

6.3.1 ARX model

$$v(t) = a_1 v(t-1) + ... + a_{n_a} v(t-n_a) + b_1 u(t-1) + ... + b_{n_b} u(t-n_b) + \eta(t) A(z) v(t) \quad = B(z) u(t-1) +$$

Transfer functions from u and η

$$W_u(z) = \frac{B(z)}{A(z)} \qquad\qquad W_\eta(z) = \frac{1}{A(z)}$$

6.3.2 ARMAX model

$$\begin{aligned} A(z)v(t) &= C(z)\eta(t) + B(z)u(t-1) \\ y(t) &= W(z)\eta(t) + G(z)u(t) \end{aligned}$$

Predictor

$$\hat{y}(t|t-1) = \frac{C(z) - A(z)}{C(z)} y(t) + \frac{B(z)}{C(z)} u(t-1)$$

Part II
Identification

Consists of estimating a model from data.

7 Prediction Error Minimization

Aims to minimize $\epsilon(t) = v(t) - \hat{v}(t|t-r)$
Steps:

1. **Data collection:** collect $\vec{u}$ and $\vec{y}$
2. **Family selection:** choose a family of models $M(\theta)$

   MA(1) $\theta = [a]$

   MA(n) $\theta = [a_1, ..., a_n]$

   ARMA(n_a, n_c) $\theta = [a_1, ..., a_{n_a}, c_1, ..., c_{n_c}]$

   ...

3. **Select an optimization criterion**

   Mean Squared error $J(\theta) = \frac{1}{N}\sum_{t=1}^{N} \epsilon_\theta(t)^2$

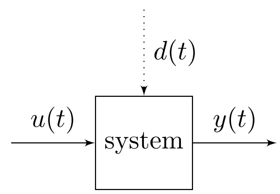
   Mean absolute error $J(\theta) = \frac{1}{N}\sum_{t=1}^{N} |\epsilon_\theta(t)|$

   ...

4. **Optimization** find $\hat{\theta} = argmin\, J(\theta) \implies \frac{dJ(\theta)}{d\theta} = 0$
5. **Validation** verify if the result satisfies the requirements

Part III
Black-Box non-parametric I/O systems

A State-space models



Known (measured) data

$\{u(1), \ldots, u(N)\}$	input
$\{y(1), \ldots, y(N)\}$	output

A.1 State-space representation

$$\begin{cases} x(t+1) = Fx(t) + Gu(t) & \text{state equations} \\ y(t) = Hx(t) + Du(t) & \text{output equations} \end{cases}$$

Where $F_{n \times n}$, $G_{n \times 1}$, $H_{1 \times n}$ and $D_{1 \times 1}$ are matrices.

S.S. representation is not unique Given any invertible matrix T, let $F_1 = TFT^{-1}$, $G_1 = TG$, $H_1 = HT^{-1}$, $D_1 = D$. Then the system $\{F, G, H, D\}$ is equivalent to the system $\{F_1, G_1, H_1, D_1\}$.

A.2 Transfer function representation

$$W(z) = \frac{B(z)}{A(z)} z^{-k} = \frac{b_0 + b_1 z^{-1} + \cdots + b_p z^{-p}}{a_0 + a_1 z^{-1} + \cdots + a_n z^{-n}} z^{-k}$$

$W(z)$ is a rational function of the z operator $\rightarrow$ is a digital filter

Infinite impulse response $W(z) = \frac{z^{-1}}{1 + \frac{1}{3} z^{-1}}$

Finite impulse response $W(z) = z^{-1} + \frac{1}{2} z^{-2} + \frac{1}{4} z^{-3}$

A.3 Convolution of the input with the inpulse response

Let's call $\omega(1), \omega(2), \ldots$ the values of $y(t)$ when $u(t) = \text{impulse}(0)$, and let's measure the values of y at different times: . Then it can be proven that for any $u(t)$

$$y(t) = \sum_{k=0}^{\infty} \omega(k) u(t-k)$$

B Converting representations one to another

B.1 State space to Transfer function

Consider a strictly propter system:

$$\begin{cases} x(t+1) = Fx(t) + Gu(t) \\ y(t+1) = Hx(t) + \cancel{Du(t)}^{\,0} \end{cases} \Rightarrow \begin{cases} x(t+1) = Fx(t) + Gu(t) \\ y(t) = Hx(t) \end{cases}$$

Applying the z operator we get

$$\begin{aligned} zx(t) &= Fx(t) + Gu(t) \\ x(t)(zI - F) &= Gu(t) \\ x(t) &= (zI - F)^{-1}Gu(t) \\ y(t) &= H(zI - F)^{-1}Gu(t) \end{aligned}$$

And we can extract the transfer function:

$$W(z) = H(zI - F)^{-1}G$$

B.2 Transfer Function to State Space

We have the transfer function

$$W(z) = \frac{b_0 z^{n-1} + b_1 z^{n-2} + \cdots + b_{n-1}}{z^n + a_0 z^{n-1} + \cdots + a_n}$$

The formulas for the state space matrices is

$$F = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \\ -a_n & -a_{n-1} & \dots & \dots & -a_1 \end{bmatrix} \quad G = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad H = \begin{bmatrix} b_{n-1} & b_{n-2} & \dots & b_0 \end{bmatrix} \quad D = 0$$

B.3 Transfer function to Impulse response

Obtained by computing the ∞ long division of $W(z)$

B.4 Impulse response to Tranfer function

Z-transform Given a discrete-time signal $s(t)$ such that $\forall t < 0 : s(t) = 0$, it's Z-transform is

$$\mathcal{Z} = \sum_{t=0}^{\infty} s(t) z^{-t}$$

It can be proven that:

$$W(z) = \mathcal{Z}(\omega(t)) = \sum_{t=0}^{\infty} \omega(t) z^{-1}$$

NB: this works only in theory because of the infinite sum

B.5 State space to Impulse response

Consider the state space model:

$$\begin{cases} x(t+1) = Fx(t) + Gu(t) \\ y(t) = Hx(t) \end{cases}$$

We have that:

$$\begin{aligned}
x(1) &= \cancelto{0}{Fx(0)} + Gu(0) & &= Gu(0) \\
y(1) &= Hx(1) & &= HGu(0) \\
\\
x(2) &= Fx(1) + Gu(1) & &= FGu(0) + Gu(1) \\
y(2) &= Hx(2) & &= HFGu(0) + HG(u1) \\
\\
x(3) &= Fx(2) + Gu(2) & &= F^2Gu(0) + FGu(1) + Gu(2) \\
y(3) &= Hx(3) & &= HF^2Gu(0) + HFGu(1) + HGu(2) \\
&\vdots
\end{aligned}$$

$$y(t) = 0u(t) + HGu(t-1) + HFGu(t-2) + HF^2Gu(t-3) + \dots$$

The IR is:

$$\omega(t) = \begin{cases} 0 & \text{if } t = 0 \\ HF^{t-1}G & \text{if } t > 0 \end{cases}$$

C Controllability and Observability

$$\begin{cases} x(t+1) = Fx(t) + Gu(t) \\ y(t) = Hx(t) \end{cases}$$

Fully observable system The system is fully observable (from the output) $\iff$ the observability matrix is full rank:

$$O = \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix} \qquad rank(O) = n$$

Fully controllable system The system is fully controllable (from the input) $\iff$ the controllability (also called reachability) matrix is full rank:

$$R = \begin{bmatrix} G & FG & \dots & F^{n-1}G \end{bmatrix} \qquad rank(R) = n$$

D Hankel Matrix

Starting from $\omega(1), \omega(2), \ldots, \omega(N)$ where $N \geq 2n-1$, we can build the Hankel Matrix of order n:

$$H_n = \begin{bmatrix} \omega(1) & \omega(2) & \omega(3) & \ldots & \omega(n) \\ \omega(2) & \omega(3) & \omega(4) & \ldots & \omega(n+1) \\ \omega(3) & \omega(4) & \omega(5) & \ldots & \omega(n+2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega(n) & \omega(n+1) & \omega(n+2) & \ldots & \omega(2n-1) \end{bmatrix}$$

Knowing that

$$\omega(t) = \begin{cases} 0 & \text{if } t = 0 \\ HF^{t-1}G & \text{if } t > 0 \end{cases}$$

We can rewrite

$$H_n = \begin{bmatrix} HG & HFG & HF^2G & \ldots & HF^{n-1}G \\ \vdots & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ HF^{n-1}G & \ldots & \ldots & \ldots & HF^{2n-2}G \end{bmatrix} = \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix} \cdot \begin{bmatrix} G & FG & \ldots & F^{n-1}G \end{bmatrix} = O \cdot R$$

E Subspace-based State Space System Identification

Impulse experiment Measure $y(t)$ under the input $u(t) = \text{impulse}(0)(0)$
How to derive F, G, H from $\omega(0), \dots, \omega(n)$?

- Assuming the IR measurement to be noise free $\rightarrow$ easier, not realistic
- Measure $\hat{\omega(t)}$ as a noisy signal and compute $\omega(t) = \eta(t) - \hat{\omega}(t)$

E.1 Obtain F, G, H from a noise-free IR

1. Build the Hankel matrix of increasing order, and conpute the rank until $rank(H_n) = rank(H_{n+1})$. Then, n is the order of the IR

$$H_1 = \begin{bmatrix}\omega(1)\end{bmatrix} \quad H_2 = \begin{bmatrix}\omega(1) & \omega(2) \\ \omega(2) & \omega(3)\end{bmatrix} \quad H_3 = \dots \quad \dots \quad H_n = \dots$$

2. Take H_{n+1} and factorize it in two rectangular matrix of size $(n+1) \times n$ and $n \times (n+1)$: $H_{n+1} = O_{n+1} \cdot R_{n+1}$, where

$$O_{n+1} = \begin{bmatrix} H \\ HF \\ \vdots \\ HF^n \end{bmatrix} \qquad R_{n+1} = \begin{bmatrix} G & FG & \dots & F^nG \end{bmatrix}$$

3. Estimate H,F,G:
   - Extract F and G from the first element of O and R
   - Define:

$$O_1 = \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix} \qquad O_2 = \begin{bmatrix} HF \\ \vdots \\ HF^n \end{bmatrix}$$

   - Observe that $O_1 F = O_2$, so $F = O_1^{-1} O_2$

F Obtain F, G, H from a noisy IR

The measurement is of $\hat{\omega}(t) = \omega(t) + \eta(t)$. To identify the process:

1. Build the Hankel matrix from data using all the N data available in one shot:

$$\hat{H}_{q\times d} = \begin{bmatrix} \hat{\omega}(1) & \hat{\omega}(2) & \dots & \hat{\omega}(d) \\ \hat{\omega}(2) & \hat{\omega}(3) & \dots & \hat{\omega}(d+1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\omega}(q) & \hat{\omega}(q+1) & \dots & \hat{\omega}(q+d+1) \end{bmatrix}$$

   Where $q + d + 1 = N$

2. Calculate the Singular Value Decomposition of $\hat{H}_{q\times d}$:

$$\hat{H}_{q\times d} = \hat{U}_{q\times q} \cdot \hat{S}_{q\times d} \cdot \hat{V}^T_{d\times d}$$

   $\hat{U}$ and $\hat{V}$ are unitary matrices: they are invertible and their inverses are equal to their transpose.

$$\hat{S} = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_d \end{bmatrix}$$

3. Plot the singular values (σ_i)and cut-off the three matrices:
   - Ideally, after a certain n (the order of the IR) there would be a jump dividing the signal (before) from the noise (after)
   - In reality no clear distinction exists, but it's possible to identify an interval of possible values of n. A tradeoff between complexity, precision and oferfitting takes place

4. Split $\hat{U}, \hat{S}, \hat{V}^T$ obtaining $U_{q\times n}$, $S_{n\times n}$, $V^T_{n\times d}$ and then recreate $H_{qd} = USV^T$

5. H and G are estimated as for the unnoisy case. To estimate F we can build O_1 and O_2 as before, but then the system $O_1 \cdot F = O_2$ cannot be solved directly as O_1 is not square. We can instead compute the approximate least-square solution of the system:

$$F = (O_1^T O_1)^{-1} O_1^T O_2$$

Part IV
Cheatsheet

G Probability Recall

Cross-Variance $Var[v,u] = E[(v - E[v])(u - E[u])]$

Variance Matrix $\begin{vmatrix} Var[v_1] & . & . & Var[v_1, v_k] \\ . & . & & . \\ . & & . & . \\ Var[v_k, v_1] & . & . & Var[v_k] \end{vmatrix}$

Covariance coefficient $\delta[i,j] = \frac{Var[i,j]}{\sqrt{Var[i]}\sqrt{Var[j]}}$

Stationary process

- m constant
- λ^2 constant
- covariance $\gamma(\tau)$ depends only on time difference
- $|\gamma(\tau)| \leq \gamma(0) \quad \forall \tau$

White noise $\eta(t) \sim WN(m, \lambda^2)$

- Stationary process
- $\gamma(\tau) = 0 \quad \forall \tau \neq 0$
- $v(t) = \alpha\eta(t) + \beta \implies v(t) \sim WN(\beta, \alpha^2\lambda^2)$

Canonical representation

- Monic
- Same degree
- Coprime
- Poles and zeros in unit disk

H Spectral analysis

Spectrum

- $\Gamma(\omega) = \gamma(0) + 2cos(\omega)\gamma(1) + 2cos(2\omega)\gamma(2) + ...$
- Periodic $T = 2\pi$
- Even
- $\Gamma_\eta(\omega) = \gamma_\eta(0) = \lambda^2$

Complex spectrum

- $\Phi(z) = \sum_{\tau=-\infty}^{+\infty} \omega(\tau) z^{-\tau}$
- $\Gamma(\omega) = \Phi(e^{j\omega})$

Fundamental theorem of spectral analysis

- $\Gamma_{\text{out}}(\omega) = |W(e^{j\omega})|^2 \cdot \Gamma_{\text{in}}(\omega)$
- $\Phi_{\text{out}}(z) = W(z)W(z^{-1}) \cdot \Phi_{\text{in}}(z)$

I Moving Average MA(n)

- $W(z) = \frac{c_0 z^n + c_1 z^{n-1} + ... + c_n}{z_n}$
- $m = 0$
- $\gamma(\tau) = \begin{cases} (c_0 c_\tau + c_1 c_{1+\tau} + ... + c_{n-\tau} c_\tau)\lambda^2 & |\tau| \leq n \\ 0 & \text{otherwise} \end{cases}$

I.1 MA(∞)

- $\gamma(0) = (c_0^2 + c_1^2 + ... + c_k^2 + ...)\lambda^2$
- $\gamma(0)$ must converge to a finite value

J Auto Regressive AR(n)

- $m = 0$
- $W(z) = \frac{z^n}{z^n - a_1 z_{n-1} - \ldots - a_n}$
- Covariance calculated by its definition

K Known predictors

AR(1) $\hat{v}(t|t-r) = a^r v(t-r)$

MA(1) $\hat{v}(t|t-1) = v(t-1) - c\hat{v}(t-1|t-2)$

MA(n) $\hat{v}(t|t-\mathbf{k}) = 0 \quad \forall k > n$

ARMA(n_a, n_b) $\hat{v}(t|t-1) = \frac{C(z)-A(z)}{C(z)} v(t)$

ARMAX(n_a, n_b) $\hat{y}(t|t-1) = \frac{C(z)-A(z)}{C(z)} y(t) + \frac{B(z)}{C(z)} u(t-1)$