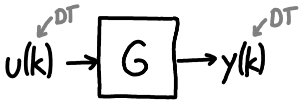
System identification

Notation differences to signals + systems:

- DT notation: $x[n]$ is now $x(k)$
- DT-fourier-transform: $X(\Omega)$ is now $X(e^{i\omega})$; $\Omega \to \omega$
- Discrete-fourier-series: $X[k]$ is now $X(e^{i\omega_n})$; $\Omega = \frac{2\pi k}{N} \to \omega_n = \frac{2\pi n}{N}$

Fundamental Concept:



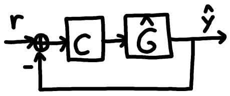
1) Choose input $u(k)$, $k=1,2,\ldots$
2) Apply $u(k)$ to plant G
3) Measure output $y(k)$, $k=1,2,\ldots$
4) Get model of G from $u(k), y(k)$ $\Rightarrow$ $y = G \cdot u \to \hat{G} \approx y/u$

- **frequency domain** (DT-Fourier Transform)
  Estimate $U(e^{i\omega}), Y(e^{i\omega}) \to$ get $\hat{G}(e^{i\omega})$
  (requires more data, has less noise problems)
- **time domain** (Z-Transform)
  get $\hat{G}(z)$ directly from $u(k), y(k)$
  (requires less data, has more noise problems)

- **open-loop**: $\hat{y} = \hat{G} u$

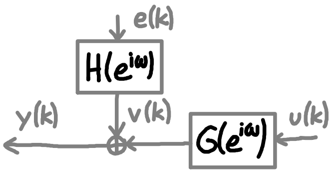
  $\hookrightarrow \|y - \hat{y}\|$ low!

- **closed-loop**: $\hat{y} = \frac{\hat{G}C}{1+\hat{G}C} \cdot r$ (if G unstable)



  $\hookrightarrow \frac{\hat{G}C}{1+\hat{G}C}$ stable!, $\left\| \frac{\hat{G}C}{1+\hat{G}C} - \frac{\hat{G}C}{1+\hat{G}C} \right\|$ low!

Open-loop configuration



G: plant to identify
H: noise filter
e: white noise
v: filtered noise
u: input
y: output

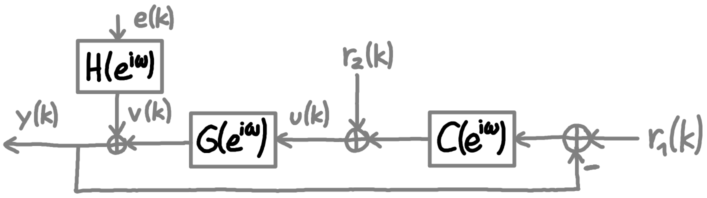
real: $y(k) = \sum_{l=0}^{\infty} g(l) u(k-l) + v(k)$

assumed: $y(k) = \sum_{l=0}^{N-1} g(l) u(k-l)$

$Y(e^{i\omega_n}) = \hat{G}(e^{i\omega_n}) \cdot U(e^{i\omega_n})$

- u periodic $\to$ no transient in y
- N with many periods of u $\to$ lower noise effect
- many samples per period $\to$ higher resolution $\hat{G}$

Closed-loop configuration



r_1: reference
r_2: excitation
$\Big\}$ $r = r_2 + C r_1$ $\leftarrow$ $\boxed{\text{different from usual } r = r_1,\ r_2 = 0\ !}$

$y = SGr + Sv$
$u = Sr - SCv$

$S = 1/(1+CG)$: sensitivity

Quality of model estimation $\hat{G}(e^{i\omega})$

- Bias: $\text{Bias}(\hat{G}) = G - E\{\hat{G}\}$
- Variance: $\text{var}(\hat{G}) = E\{|\hat{G} - E\{\hat{G}\}|^2\}$
- Mean-square error: $\text{MSE}(\hat{G}) = E\{|G - \hat{G}|^2\}$

$\hat{G}$: many estimations of plant (all different due to noise)
$E\{[\,]\}$: expected value = mean
$\to \text{MSE}(\hat{G}) = \text{var}(\hat{G}) + \text{Bias}(\hat{G})^2$

- coherency spectrum: $\hat{\kappa}_{yu}(e^{i\omega_n}) = \sqrt{|\hat{\Phi}_{yu}|^2 / \hat{\Phi}_y \hat{\Phi}_u}$

$\forall \omega_n$: if output per. from noise ($\kappa = 0$) or input ($\kappa = 1$)

Discrete Fourier Series (DFS)

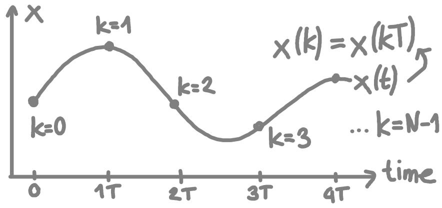
$$X(e^{i\omega_n}) = \sum_{k=0}^{N-1} x(k) e^{-i\omega_n k} \quad \begin{pmatrix} \omega_n = \omega_{in} \cdot T \\ \omega_n = 2\pi n / N \end{pmatrix}$$

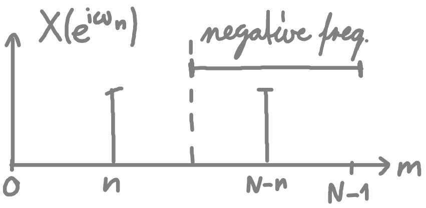
neg. freq: $X(e^{i\omega_n}) = X(e^{i\omega_{N-n}})$

amplitude: $\begin{cases} n < \frac{N}{2}: & a_n = \frac{2}{N} \cdot \|X(e^{i\omega_n})\| \\ n = \frac{N}{2}: & a_{N/2} = \frac{1}{N} \|X(e^{i\omega_{N/2}})\| \\ n = 0: & a_0 = \frac{1}{N} \cdot \|X(e^{i\omega_0})\| \end{cases}$

offset: ($n=0$ case above)

phase: $\varphi_n = \measuredangle(X(e^{i\omega_n}))$





T: sampling time
x(t): continuous signal
x(k): sampled signal
k: sample # from 0 to M
N: # of samples k measured
τ_p: meas. time ($\tau_p = TN$)

$$\boxed{\text{eg. } x(t) = \sin(4\cdot 2\pi \cdot t) \quad N = 120 \quad T = 1/30 \quad \rightarrow \omega_n = 4\cdot 2\pi \cdot \tfrac{1}{30} = 2\pi \tfrac{n}{120} \rightarrow n = 16 \rightarrow \{ \text{peak at } n = 16}$$

! DFS assumes signal repeats every N. If it does not (eg. $x(k)=0\ \forall k<0$) transient error is introduced !

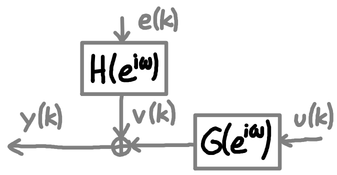
Other fourier function properties

$x(k), k = 0, \dots, N-1 \rightarrow X(e^{i\omega_n}),\ \omega_n = \frac{2\pi n}{N},\ n = 0, \dots, N/2$

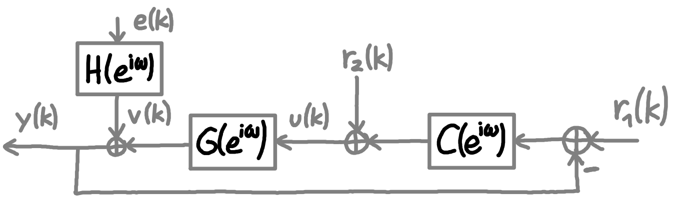
		finite discrete (assumes periodic)	infinite discrete	continuous
transform:	$X(e^{i\omega_n}) =$	$\sum_{k=0}^{N-1} x(k) e^{-i\omega_n k}$	$\sum_{k=-\infty}^{\infty} x(k) e^{-i\omega k}$	$\int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt$
		$\frac{1}{N} \sum_{k=0}^{N-1} X(e^{i\omega_n}) e^{i\omega_n k}$	$\frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{i\omega}) e^{i\omega k} d\omega$	$\int_{-\infty}^{\infty} X(\omega) e^{i\omega t} d\omega$
autocorrelation:	$R_x(\tau) =$	$\frac{1}{N} \sum_{k=0}^{N-1} x(k) \cdot x(k-\tau)$	$\sum_{k=-\infty}^{\infty} x(k) \cdot x(k-\tau)$	?
spectral dens.:	$\Phi_x(e^{i\omega_n}) =$	$\frac{1}{N} \|X(e^{i\omega})\|^2 = \mathcal{F}(R_x(\tau))$	$\|X(e^{i\omega})\|^2 = \mathcal{F}(R_x(\tau))$	?
energy:	$E =$	$\sum_{k=0}^{N-1} \|x(k)\|^2 = \sum_{n=0}^{N-1} \Phi_x(e^{i\omega_n})$	?	?
cross-correlation:	$R_{xy}(\tau) =$	$\frac{1}{N} \sum_{k=0}^{N-1} x(k) y(k-\tau)$	?	?
cross-spectral:	$\Phi_{xy}(e^{i\omega_n}) =$	$\frac{1}{N} X(e^{i\omega_n}) \cdot Y^*(e^{i\omega_n}) = \mathcal{F}(R_{xy}(\tau))$	?	?

$R_x(-\tau) = R_x(\tau)$
$R_x(0) \geq R_x(\tau) \quad \forall \tau > 0$
$\Phi_{xy}(e^{i\omega}) = \Phi_{yx}(e^{i\omega})$

$\Phi_x(e^{i\omega}) \in \mathbb{R}$
$\Phi_x(e^{i\omega}) \geq 0 \quad \forall \omega$
$\Phi_x(e^{i\omega}) = \Phi_v(e^{-i\omega})$



$\Phi_y = |G|^2 \cdot \Phi_u$
$\Phi_{yu} = G \cdot \Phi_u$
$\Phi_v = \Phi_y - |\Phi_{uy}|^2 / \Phi_u$



$r = r_2 + C r_1$
$S = 1/(1+CG)$
$T = CG/(1+CG)$
$L = CG$

$\Phi_{yu} = |S|^2 G \Phi_r - |S|^2 C^* \Phi_v$
$\Phi_u = |S|^2 \Phi_r + |S|^2 |C|^2 \Phi_v$

$y = SGr + Sv \leftarrow$
$u = Sr - SCv \leftarrow$

unusual r! with $r = r_1, r_2 = 0$:
$y = Tr + Sv$
$u = SCr - SCv$

Input signals

- **steps/doublets**:

$u(k)=\begin{cases} k\geq 0 : A \\ \text{else} : 0 \end{cases} \rightarrow U(e^{i\omega})=1/i\omega$

$u(k)=\begin{cases} 0\geq k\geq \varepsilon/2 : A \\ \varepsilon/2\geq k\geq \varepsilon : -A \\ \text{else} : 0 \end{cases} \rightarrow U(e^{i\omega})= \dots$

- **sinusoids**:

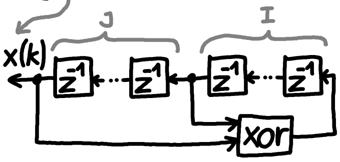
$u_n(k)=\sin(\omega_n k) \rightarrow U(e^{i\omega})=\text{peak}@\omega_n$

- **filtered white noise**:

$U(e^{i\omega})=\underbrace{L(e^{i\omega})}_{\text{filter}}\cdot \underbrace{e}_{\text{white noise}}$

- **PRBS**: (periodic with $2^{I+J}-1$)

$u(k)=\begin{cases} x(k)=1 : +A \\ x(k)=0 : -A \end{cases} \rightarrow |U(e^{i\omega})| \approx \text{const.}$



I+J	J
3	1 or 2
4	1 or 3
5	2 or 3
6	1 or 5
7	1 or 6
8	—
9	4 or 5
⋮	⋮

- **multi-sinusoidal**

$u(k)=\sum_{s=1}^{S}\sqrt{2\alpha_s}\cos(\omega_s kT+\phi_s)$ $\quad (S<N/2L)$

$\underbrace{\sum_{s=1}^{S}\alpha_s=1}_{\text{sig. pow}=1}, \underbrace{\omega_s=\frac{2\pi l s}{N}\ L\in\mathbb{N}}_{\text{harmonic frequencies}}, \underbrace{\phi_s=2\pi\sum_{n=1}^{s} n\alpha_s}_{\text{minimize peak } u(k)}$

Improve model estimation

- **averaging**: divide experiment data into R parts and estimate $\hat{G}_r(e^{i\omega_n})$ of each → average to $\tilde{G}(e^{i\omega_n})$

$\tilde{G}(e^{i\omega_n})=\sum_{r=1}^{R}\alpha_r\cdot\hat{G}_r(e^{i\omega_n})$ $\quad$ (average : $\alpha_r=1/R$ $\quad$ min variance: $\alpha_r(e^{i\omega_n})=|U_r(e^{i\omega_n})|^2/\sum_{r=1}^{R}|U_r(e^{i\omega_n})|^2$)

↳ $R\uparrow \Rightarrow \text{var}(\tilde{G})\downarrow$, $\text{bias}(\tilde{G})\uparrow$ ⇒ tradeoff! find R with lowest $\text{MSE}(\tilde{G})$ (with ref. dataset!)

(due to transient in each r. not the case if u periodic)

- **smoothing**:

$\tilde{G}(e^{i\omega_n})=\frac{1}{2\pi}\int_{-\pi}^{\pi}W_\gamma(e^{i(\xi-\omega_n)})\frac{1}{N}|U_N(e^{i\xi})|^2\hat{G}(e^{i\xi})d\xi \Big/ \frac{1}{2\pi}\int_{-\pi}^{\pi}W_\gamma(e^{i(\xi-\omega_n)})\frac{1}{N}|U(e^{i\xi})|^2 d\xi$

$\approx \sum_{k=0}^{N}W_\gamma(e^{i\omega_{k-n}})\cdot|U(e^{i\omega_k})|^2\cdot\hat{G}(e^{i\omega_k}) \Big/ \sum_{k=0}^{N}W_\gamma(e^{i\omega_{k-n}})\cdot|U(e^{i\omega_k})|^2$

$\approx \sum_{\tau=-\gamma}^{\gamma}w_\gamma(\tau)\hat{R}_{yu}(\tau)e^{-i\tau\omega_n} \Big/ \sum_{\tau=-\gamma}^{\gamma}w_\gamma(\tau)\hat{R}_u(\tau)e^{-i\tau\omega_n}$

$W_\gamma(e^{i\omega})$: FD window function
$W_\gamma(e^{i\omega_n})$: DFD window function
$w_\gamma(\tau)$: TD window function

↳ $\gamma\downarrow \Rightarrow \text{var}(\tilde{G})\downarrow$, $\text{bias}(\tilde{G})\uparrow$ ⇒ tradeoff! find γ with lowest $\text{MSE}(\tilde{G})$ (with ref. dataset!)

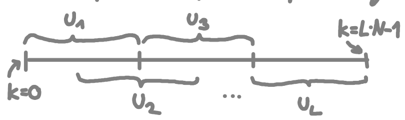
e.g. Bartlett window: $w_\gamma(\tau)=1-|\tau|/\gamma$, $-\gamma\leq\tau\leq\gamma \leftrightarrow W_\gamma(e^{i\omega})=\frac{1}{\gamma}\left(\frac{\sin\gamma\omega/2}{\sin\omega/2}\right)^2$ $\quad$ or Hann, Hamming, Welch,...

- **TD windowing**:

$U_w(e^{i\omega_n})=\sum_{k=0}^{N-1}w_{data}(k)\cdot u(k)\cdot e^{-ik\omega_n}$ $\quad$ ($w_{data}(k)=w_\gamma(k-N/2)|_{\gamma=N/2}$: some window over whole N centered at N/2)

$E_{scl}=\sum_{k=0}^{N-1}|w_{data}(k)u(k)|^2 \Big/ \sum_{k=0}^{N-1}|u(k)|^2$ $\quad$ (scale factor to preserve pow. in periodigram $\Phi_{u_w}=\frac{1}{E_{scl}}\cdot\frac{1}{N}\cdot|U_w(e^{i\omega_n})|^2$)

Welch's method: $\Phi_u(e^{i\omega_n})=\frac{1}{NL}\sum_{l=1}^{L}\frac{1}{E_{scl}}|U_l(e^{i\omega_n})|^2$



$U_l=U_w$ for $u=u_l$ with $l=1\dots L$

→ advantages: reduce transient effect $\quad$ → disadvantages: "smears" frequencies

- **TD detrending**: $u_d(k)=u(k)-(\alpha k+\beta)$ $\quad$ ($\alpha k+\beta$: best linear fit) (only do this when drift/offset unwanted)

Frequency-Domain open-loop identification

- **Sinusoidal correlation:** (sine test signal with $\omega_v \to$ measure output phase+amplitude $\to$ reconstruct $\hat{G}$)

input: $u(k) = \alpha \cos(\omega_v k)$ $\quad (k = 0, \dots, N-1)$

output: $y(k) = \alpha \underbrace{|G(e^{i\omega_v})|}_{\text{unknowns}} \cdot \cos(\omega_v k + \underbrace{\measuredangle(G(e^{i\omega_v}))}_{\text{unknowns}}) + \underbrace{\text{noise } v(k) + \text{transient}}_{\text{disturbances}}$

- record many periods
- start when transient gone

correlation functions

$$I_c = \frac{1}{N}\sum_{k=0}^{N-1} y(k)\cos(\omega_v k)$$
$$I_s = \frac{1}{N}\sum_{k=0}^{N-1} y(k)\sin(\omega_v k)$$

$\Rightarrow$ expected val. for $N \to \infty$, exp. variance for $N \to \infty$:

$$E\{I_c\} = \frac{\alpha}{2}|G(e^{i\omega_v})|\cos(\measuredangle G(e^{i\omega_v})), \quad \text{var}\{I_c\} = 0$$
$$E\{I_s\} = \frac{-\alpha}{2}|G(e^{i\omega_v})|\sin(\measuredangle G(e^{i\omega_v})), \quad \text{var}\{I_s\} = 0$$

$$\hookrightarrow \hat{G}(e^{i\omega_v}) = \frac{I_c^2 + I_s^2}{\alpha/2}, \quad |G(e^{i\omega_v})| = \frac{\sqrt{I_c^2 + I_s^2}}{\alpha/2} \quad \measuredangle G(e^{i\omega_v}) = -\arctan\left({}^{I_s}/_{I_c}\right)$$

- **Empirical transfer function estimation (ETFE):**

input $u(k)$: **anything** $\to$ fourier: $U(e^{i\omega_n})$
output $y(k)$: **measured** $\to$ fourier: $Y(e^{i\omega_n})$

$$\Big\} \ \hat{G}(e^{i\omega_n}) = Y(e^{i\omega_n}) / U(e^{i\omega_n}) \qquad (\text{real: } Y/U = G + V/U)$$

best results if
- high signal/noise ratio $(U \gg V)$ $\Rightarrow$ $u(k)$ as large as possible; take periodic $u(k)$ and record multiple periods!
- no transient / periodic signal U $\Rightarrow$ ignore first couple of periods

- **Spectral estimation:**

opt. 1
input $u(k)$: **anything** $\to$ fourier: $U(e^{i\omega_n})$
output $y(k)$: **measured** $\to$ fourier: $Y(e^{i\omega_n})$

$$\to \Phi_{uu}(e^{i\omega_n}) = \frac{1}{N}|U(e^{i\omega_n})|^2, \quad \to \Phi_{uy}(e^{i\omega_n}) = \frac{1}{N}U(e^{i\omega_n})\cdot Y^*(e^{i\omega_n}) \quad \Big\} \ \hat{G}(e^{i\omega_n}) = \Phi_{yu}(e^{i\omega_n}) / \Phi_{uu}(e^{i\omega_n})$$

opt. 2
input $u(k)$: **anything**
output $y(k)$: **measured**

$$\to R_{uu}(\tau) = \frac{1}{N}\sum_{k=0}^{N-1} u(k)u(k-\tau) \ \to \text{fourier: } \Phi_{uu}(e^{i\omega_n})$$
$$\to R_{uy}(\tau) = \frac{1}{N}\sum_{k=0}^{N-1} u(k)y(k-\tau) \ \to \text{fourier: } \Phi_{uy}(e^{i\omega_n})$$
$$\Big\} \ \hat{G}(e^{i\omega_n}) = \Phi_{yu}(e^{i\omega_n}) / \Phi_{uu}(e^{i\omega_n})$$

! same best result tips as for ETFE! $\quad$ ↑ alt. formula without assumption of periodicity?

FD subspace identification

(get state-space $\hat{A}, \hat{B}, \hat{C}, \hat{D}$ from $\hat{G}(e^{i\omega_n}) = \hat{G}(n)$)

$\hat{G}(e^{i\omega_n})$ (e.g. from ETFE)

$\downarrow$

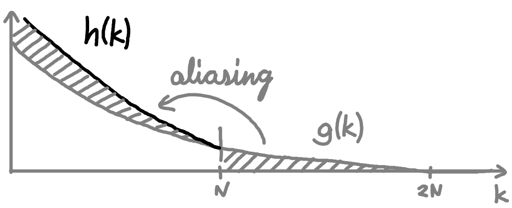
$$x(k+1) = \hat{A}x(k) + \hat{B}u(k)$$
$$y(k) = \hat{C}x(k) + \hat{D}u(k)$$

impulse resp.: $g(k) = \begin{cases} 0 & k<0 \\ \hat{D} & k=0 \\ CA^{k-1}\cdot B & k>0 \end{cases} = \mathcal{F}^{-1}(G(e^{i\omega}))$ (continuous FT!)

aliased imp. resp.: $h(k) = CA^{k-1}\left(\sum_{l=0}^{N} A^{Nl}\right)B = \mathcal{F}^{-1}(G(e^{i\omega_n})) \approx \mathcal{F}^{-1}(\hat{G}(e^{i\omega_n}))$ (discrete FT!)

non trivial, because $\mathcal{F}^{-1}(\hat{G}(e^{i\omega_n})) \neq g(k)$ impulse response. actually aliased impulse resp. $h(k)$

$\hookrightarrow$ solution: subspace identification algorithm (lecture 7)
(+ alternative for when ω_n non-uniformly spaced)



Frequency-Domain close-loop identification (estimate $\hat{G}$ while it is being controlled by C)

usefull when: plant unstable or need to stay at operating point.

! watch out, $r \neq$ usual r !

- **direct method**: (like for open-loop ETFE and spectral estimate)

(worked well in open-loop because $\Phi_{uv}=0$, but in closed loop $\Phi_{uv}\neq 0$ so worse estimate!)

$$\hat{G}(e^{i\omega_n}) \approx Y(e^{i\omega_n})/U(e^{i\omega_n}) = \frac{SGr+Sv}{Sr+SCv} \qquad \hat{G}(e^{i\omega_n}) \approx \Phi_{yu}(e^{i\omega_n})/\Phi_{uu}(e^{i\omega_n}) = \frac{|S|^2 G\Phi_r - |S|^2 C^* \Phi_v}{|S|^2\Phi_r + |S|^4 |C|^2 \Phi_v} \qquad (S = 1/1+CG)$$

only works if: S not too low (if C is too good it "hides" G) ; $r \gg v$ resp. $\Phi_r \gg \Phi_v$ (more signal than noise)

- **indirect method**: $\frac{\hat{Y}(e^{i\omega_n})}{\hat{R}(e^{i\omega_n})} = \hat{T}_{yr}(e^{i\omega_n}) = \frac{\hat{G}}{1+C\hat{G}} \rightarrow \hat{G}(e^{i\omega_n}) = \frac{\hat{T}_{yr}(e^{i\omega_n})}{1-\hat{T}_{yr}(e^{i\omega_n})C(e^{i\omega_n})}$

(why better?)

- **input-output method**: $\frac{\hat{Y}(e^{i\omega_n})}{\hat{R}(e^{i\omega_n})} = \hat{T}_{yr}(e^{i\omega_n}) = \frac{\hat{G}}{1+C\hat{G}}$; $\frac{\hat{U}(e^{i\omega_n})}{\hat{R}(e^{i\omega_n})} = \hat{T}_{ur}(e^{i\omega_n}) = \frac{1}{1+C\hat{G}} \rightarrow \hat{G}(e^{i\omega_n}) = \frac{T_{yr}(e^{i\omega_n})}{T_{ur}(e^{i\omega_n})}$

(why better?)

- **Youla method**: in the above methods a $\hat{G}$ can result that is not stabilized by C! (bad, since real G is)
  → use dual-Youla method to force $\hat{G}$ to be stabilized by C. (lecture 8)

Time-Domain open-loop identification

choose model structure + parameter θ ⟶ choose cost function $J(\theta)$

pulse response:	$g(k)$	$\theta=(g(0), g(1), \dots)$
transfer function:	$G(z)=\frac{b_1 z^{-1}+b_2 z^{-2}+\dots+b_m z^{-m}}{1+a_1 z^{-1}+\dots+a_n z^{-n}}$	$\theta=(a_1,\dots,a_n,b_1,\dots,b_m)$
state space:	$x(k+1)=Ax(k)+Bu(k)$ $y(k)=Cx(k)+Du(k)$	$\theta=(A_{ij},B_{ij},C_{ij},D_{ij})$

or any other norm

residual error: $J(\theta)=\|y-Gu\|_2^2, \|y-Gu\|_\infty, \|y-Gu\|_1$

parametric error: $J(\theta)=\|\theta-\theta_0\|$ (θ_0: ??)

prediction error: $J(\theta)=|y(k+1)-\hat{y}(k+1)|^2$ prediction
meas. outside experiment

⟶ solve optimization problem: minimize $J(\theta) \triangleq \hat{\theta}=\text{argmin}(J(\theta))$ (usually numerical solver)

- **correlation-based method** (identify first M entries of g(k) with N measurements of u(k), y(k) and norm2 residual error)

$R_{yu}(\tau)=g(k) * R_u(\tau)$ (accurate if u,y: periodic+noise free + $g(k)=0\ \forall k\geq M$)

$$\begin{pmatrix}\hat{R}_{yu}(0)\\ \vdots \\ \hat{R}_{yu}(N-1)\end{pmatrix} = \begin{pmatrix}\hat{R}_u(0) & \cdots & \hat{R}_u(-(M-1))\\ \vdots & & \vdots \\ \hat{R}_u(N-1) & \cdots & \hat{R}_u(?)\end{pmatrix} \cdot \begin{pmatrix}\hat{g}(0)\\ \vdots \\ \hat{g}(M-1)\end{pmatrix} \rightarrow \hat{g}=\boxed{\bar{R}}^{\boxed{+}}\cdot\hat{R}_{yu}$$

pseudoinverse: minimize $\|\bar{R}\hat{g}-\hat{R}_{yu}\|_2$

only works if u(k) is: <u>persistently exciting of order M</u>

general u(k): $\Phi_u(e^{i\omega_n})\neq 0$ for at least M frequencies ω_n.	
step func.: $M\leq 1$	PRBS: $M\leq$ period of PRBS
multi-sinusoids: $M\leq$ # of sin/cos with $0<\omega_s<\pi$	

M=n+m here

- **linear regression method** (identify G(z) with prediction error)

$$y(k)=G(z)u(k) = -a_1 y(k-1) - \dots - a_n y(k-n) + b_1 u(k-1) + \dots + b_m u(k-m)$$
$$= \varphi(k)^T \cdot \theta = [-y(k-1),\dots,-y(k-n),u(k-1),\dots,u(k-m)]\cdot[a_1,\dots,a_n,b_1,\dots,b_m]^T$$

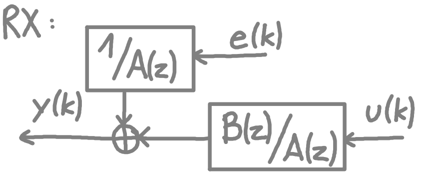
$$\rightarrow \underbrace{\begin{pmatrix}y(0)\\ \vdots \\ y(N-1)\end{pmatrix}}_{Y} = \underbrace{\begin{pmatrix}\varphi(0)^T\\ \vdots \\ \varphi(N-1)^T\end{pmatrix}}_{\Phi}\cdot\hat{\theta} \rightarrow \hat{\theta}=\Phi^{\boxed{+}}\cdot Y$$

pseudoinverse: minimize $\|\Phi\theta-Y\|_2$

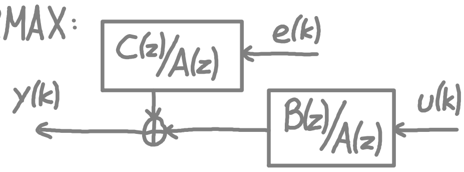
only works if u(k) is: <u>persistently exciting of order n+m</u>
and y,u noise free...

→ no bias only if ARX model ; low variance in θ with long measurement

- ARX:



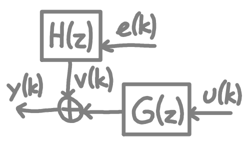
- ARMAX:



ARMAX can be made to ARX if $C(z)$ known + invertible with $u_F = C^{-1}u$, $y_F = C^{-1}y$

- **pseudo linear regression + noise prediction** (identify $G(z)$ with known $H(z)$ and error prediction)

Noise prediction: (predict $y(k)$ when y,u for $0..(k-1)$ is known and $G(z)$ known and $H(z)$ known)



$G(z)$: stable
$H(z), H_{inv}(z)$: stable
$e(k)$: white noise

$$v(k) = \sum_{i=0}^{k} h(i)e(k-i) = \underbrace{e(k)}_{\text{noise}} + \underbrace{\sum_{i=1}^{k} h(i)e(k-i)}_{\text{observable } m(k-1)}$$

($H(z)$: monic $\hat{=}$ $h(0)=1$)

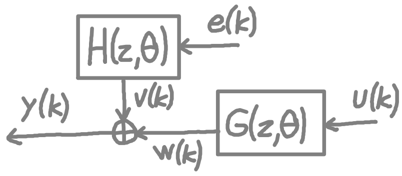
if mean(e) $\hat{=}$ most probable e $\rightarrow$ $\hat{v}(k|k-1) = m(k-1)$

$$\hat{v}(k|k-1) = m(k-1) = \ldots = (1-H_{inv}(z))v(k) \approx -\sum_{i=1}^{k} h_{inv}(i)v(k-i)$$

($H_{inv}(z) = 1/H(z)$)

$$\hat{y}(k|k-1) = G(z)u(k) + \hat{v}(k|k-1) = \ldots = H_{inv}(z)G(z)u(k) + (1-H_{inv}(z))y(k)$$

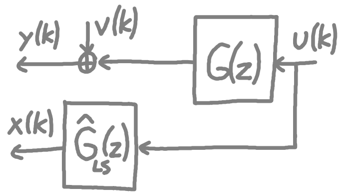
1) pick representation of $\hat{G}(z,\theta)$, $\hat{H}(z,\theta)$ ($H(z,\theta)$ must be monic $\hat{=}$ $h(0)=1$)
2) choose initial condition $\theta = \theta_0$
3) $\hat{w}(k) = \hat{G}(z)\cdot u(k)$ $\rightarrow$ $\hat{e}(k) = 1/\hat{H}(z)\cdot(y(k)-\hat{w}(k))$
4) $\hat{y}(k|k-1) = \hat{w}(k) + (\hat{H}(z)-1)\cdot\hat{e}(k)$
5) find θ that minimizes $\|\hat{e}(k)\|_2$ $\leftarrow$ (since $e(k) = y(k) - y(k|k-1)$ because $H(z)$ is monic)



no bias if real H, H^{-1} stable
lower variance in θ if measurements long

- **instrumental variable method**: (linear regression, avoiding bias)

1) estimate $G(z)$ with linear regression as if $v(k)=0\ \forall k$ $\rightarrow$ $\hat{G}_{LS}(z), \hat{\theta}_{LS}, \Phi$
2) calculate $x(k) = \hat{G}_{LS}(z)\cdot u(k)$
3) build $Z = \begin{pmatrix} \varsigma(0)^T \\ \vdots \\ \varsigma(N-1)^T \end{pmatrix}$ analogous to $\Phi = \begin{pmatrix} \varphi(0)^T \\ \vdots \\ \varphi(N-1)^T \end{pmatrix}$ but with $y(k)$ replaced by $x(k)$
4) get $\hat{\theta}_{IV} = \left(\frac{1}{N}\sum_{k=0}^{N-1}\varsigma(k)\varphi^T(k)\right)^{-1}\cdot\frac{1}{N}\sum_{k=0}^{N-1}\varsigma(k)y(k)$



improves linear regression to be unbiased with non ARX model

- **maximum likelihood method**: (identify $G(z)$ if $v(k)$ statistics are known)

1) write down formula for Z_θ
$z(k,\theta) = y(k) - G(z,\theta)\cdot u(k)$
$Z_\theta = [z(1,\theta)\ldots z(N,\theta)]^T$

2) calculate statistics of Z_θ from $v(k)$ (ν: mean vector, Σ: covariance matrix)
$z(k,\theta) = v(k) = H(z)e(k)$ (if known)
$Z_\theta \sim \mathcal{N}(\nu, \Sigma)$
↳ e.g. $v \sim \mathcal{N}(\mu, \lambda)$ $\rightarrow$ $\nu = \mu\cdot\text{ones}(\mathbb{R}^{N\times 1})$
$\Sigma = \lambda\cdot\text{diag}(\mathbb{R}^{N\times N})$

3) write down likelihood formula
$$f(\theta) = \frac{1}{(2\pi)^{N/2}\sqrt{\det\Sigma}}\cdot e^{-\frac{1}{2}(Z_\theta-\nu)^T\Sigma^{-1}(Z_\theta-\nu)}$$

4) calculate maximum likelihood $\hat{\theta}$
$$\hat{\theta} = \underset{\theta}{\arg\max}\, f(\theta) = \underset{\theta}{\arg\min}\,(Z_\theta-\nu)^T\Sigma^{-1}(Z_\theta-\nu)$$
$\leftarrow$ (could be brought to closed form)

- with a priori knowledge about $\theta \sim \mathcal{N}(\mu_\theta, \lambda_\theta)$: $f_{new}(\theta) = f_{old}(\theta)\cdot\frac{1}{\lambda_\theta\sqrt{2\pi}}\cdot e^{-\frac{1}{2}\left(\frac{\theta-\mu_\theta}{\lambda_\theta}\right)^2}$

TD subspace identification (get state-space $\hat{A},\hat{B},\hat{C},\hat{D}$ from $\hat{G}(z)$)

similar concept as for FD. (lecture 13)