1.1 Systems and Models

system
- input —— can be manipulated (可被操控)
- output —— of interest

op

不可 ⇒ 扰动

(一般)
- Dynamic 当前时刻的输出不只和当前时刻输入有关，也和之前输入有关
- Static：当前时刻的输出只和当前时刻输入有关

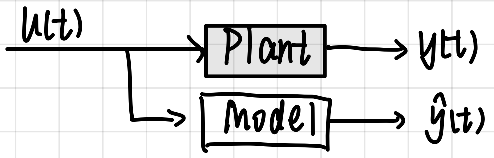
，微分方程/差分方程表征

Model Type
- Mental Model
- Graphical Model: step response, Bode, Nyquist
- Mathematical Model:
  - 连续 / 离散
  - lumped / distributed (分布式)
  - 确定 / 随机

1.2

- Mechanism Modeling (机理建模) / the first principle (第一性) . 通过物理规律建模
- system identification: from the data of input and output

1.4 Definition of system identification



The entities
- Data
- Model types
- Equivalence rule

↓ Zadeh

→
- data
- the set of models
- Criterion
- Optimization approach

1.5 Linear-parameter model

System model (常见 discrete-time) 定义 z 算子

z-transformation e.g $\{y(t)\}$, $t\in \mathbb{I}[0,+\infty)$, so $z\,y(t)=y(t+1)$

⇒ 定义: $z^{i}y(t)=y(t+i)$ (前向), $z^{-i}y(t)=y(t-i)$ (后向)

$A(z)=a_0+a_1z_1+\cdots+a_nz^n$ 有 $A(z)y(t)=a_0y(t)+a_1y(t+1)+\cdots+a_ny(t+n)$

$A(z)=B(z)\,C(z)$，则可证 $[B(z)\,C(z)]y(t)=B(z)[\,C(z)\,y(t)]$

e.g. $B(z)=b_0+b_1z$, $C(z)=c_0+c_1z$, $B(z)C(z)=b_0c_0+(b_1c_0+b_0c_1)z+b_1c_1z^2$

左边 $=b_0c_0y(t)+(b_1c_0+b_0c_1)y(t+1)+b_1c_1y(t+2)$

右边 $=B(z)\cdot(c_0y(t)+c_1y(t+1))=b_0c_0y(t)+\ldots$ = 左边

离散时间随机过程（随时间变化随机变量时间序列）

$V(t)$ $EV(t)$ white noise：① at every fixed t，$EV(t)=0$（期望）

② $E[V^2(t)]=6$，6 与 t 无关（方差）

③ $E[V(t_1)V(t_2)]=0$，$t_1\neq t_2$（协方差）

$A(z)=\sum_{i=0}^{n} a_i z^{-i}$ 取 $A(z)=1+2z^{-1}$

$\therefore A(z)V(t)=V(t)+2V(t-1)$ $\xrightarrow{验证}$ $\begin{cases} EA(z)V(t)=0 \\ E[[A(z)V(t)]^2]=56 \\ E[A(z)V(t_1)A(z)V(t_2)] \end{cases}$

$E[V(t_1)+2V(t_1-1)][V(t_2)+2V(t_2-1)]$

当 $t_2=t_1-1$，有 上式 $=26\neq 0$

$\therefore A(z)V(t)$ 不一定为白噪声

符号推广：标量 2 向量：$Y=\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$，$EY=\begin{bmatrix} EY_1 \\ \vdots \\ EY_n \end{bmatrix}$ $E[[Y-EY][Y-EY]^T]>0$

平稳随机过程：$\lim_{T\to\infty}\frac{1}{T}(\mathring{V}(0)+\mathring{V}(1)+\cdots+\mathring{V}(T))=EV(t)$

随机过程采样沿时间平均结果等于期望

数学模型：$V(t)$：white noise

时间序列（无输入）

1、moving average $y(t)=d_0V(t)+d_1V(t-1)+\cdots+d_{n_d}V(t-d_n)$

(MA)滑动平均 Z变换表达：$y(t)=(d_0+d_1z^{-1}+\cdots+d_{n_d}z^{-d_{n_d}})V(t)=D(z^{-1})V(t)$

2. autoregressive $y(t)=-c_1y(t-1)-c_2y(t-2)-\cdots-c_{n_c}y(t-n_c)+V(t)$ 需要 $y(t)$ 初值

(AR) 自回归 整理：$(1+c_1z^{-1}+\cdots+c_{n_c}z^{-n_c})y(t)=V(t) \Rightarrow C(z^{-1})y(t)=V(t)$

$y(t)=\frac{1}{C(z^{-1})}V(t)$

3、autoregressive moving average (ARMA)

$y(t)=-c_1y(t-1)-c_2y(t-2)-\cdots-c_{n_c}y(t-n_c)+V(t)+d_1V(t)+\cdots+d_{n_d}V(t-n_d)$

$C(z^{-1})=1+\sum_{i=1}^{n_c}c_iz^{-i}$ $D(z^{-1})=1+\sum_{i=1}^{n_d}d_iz^{-i} \Rightarrow C(z^{-1})y(t)=D(z^{-1})V(t)$

$y(t)=\frac{D(z^{-1})}{C(z^{-1})}V(t)$ （求解需还原成差分方程）

时间序列（有输入） $\boxed{\text{Equation error model 方程误差模型}}$

$y(t)=-a_1y(t-1)-a_2y(t-2)-\cdots-a_{n_a}y(t-n_a)+b_1u(t-1)+\cdots+b_{n_b}u(t-n_b)+w(t)$

↳ noise

$A(z^{-1})y=B(z^{-1})x+w(t)$

1、$w(t)=V(t)$

$(\underline{AR}\,\underline{X})$（有控制输入）⇒ 也称 controlled autoregressive (CAR) 受控的自回归模型

output ↗ ↳ input

定义：$A(z^{-1})=1+a_1z^{-1}+\cdots+a_{n_a}z^{-n_a}$ $\Rightarrow A(z^{-1})y$

$B(z^{-1})=b_1z^{-1}+\cdots+b_{n_b}z^{-n_b}$

2、$w(t) = v(t) + d_1 v(t-1) + \dots + d_{n_d} v(t-n_d)$ ARMAX

$D(z) = 1 + d_1 z^{-1} + \dots + d_{n_d} z^{-n_d}$

$\Rightarrow A(z^{-1}) y(t) = B(z^{-1}) u(t) + D(z^{-1}) v(t)$

3、$w(t) = \frac{1}{C(z^{-1})} v(t)$ ARARX

$\Rightarrow A(z^{-1}) y(t) = B(z^{-1}) u(t) + \frac{1}{C(z^{-1})} v(t)$

4、$w(t) = \frac{D(z^{-1})}{C(z^{-1})} v(t)$ ARARMAX ($C(z^{-1}) w(t) = D(z^{-1}) v(t)$)

$\Rightarrow A(z^{-1}) y(t) = B(z^{-1}) u(t) + \frac{D(z^{-1})}{C(z^{-1})} v(t)$

$\Longrightarrow y(t) = \frac{B(z^{-1})}{A(z^{-1})} u(t) + w(t)$ **输出误差模型**

① $w(t) = v(t)$ OE ② $w(t) = D(z^{-1}) v(t)$ OE MA

③ $w(t) = \frac{1}{C(z^{-1})} v(t)$ OE AR ④ $w(t) = \frac{D(z^{-1})}{C(z^{-1})} v(t)$ OE ARMA

How to identify a model? Recursive Q: 测量物体

最小二乘法： $J = \sum_{i=1}^{N} |l - l_i|^2$

↳ 递推最小2乘： $\underbrace{l_1, l_2, \dots, l_{t-1}}, l_t$ 有 $\left\{ \begin{array}{l} \hat{l}_{t-1} = \frac{1}{t-1}\sum_{i=1}^{t-1} l_i \\ l_t \end{array} \right.$ > $\hat{l}_t = \frac{1}{t}\sum_{i=1}^{t} l_i = \frac{1}{t}[\sum_{i=1}^{t-1} l_i + l_t] = \frac{t-1}{t}\cdot\frac{1}{t-1}\sum_{i=1}^{t-1} l_i + \frac{1}{t} l_t$

有 $\hat{l}_t = \frac{1}{N}\sum_{i=1}^{N} l_i = \frac{t-1}{t}\hat{l}_{t-1} + \frac{1}{t} l_t = \hat{l}_{t-1} + \frac{1}{t}[l_t - \hat{l}_{t-1}]$

增益 gain 新息 innovation (当前测量值与上一步估计值差)

问题：测量累积，t↗，即使 innovation 很大，估果也不变化 → data saturation

↳ model : $l(t) = L + v(t)$. suppose $v(t)$ is white noise

$\hat{\ell}(t) = \frac{1}{t}\sum_{i=1}^{t} \ell(t)$ $E\hat{\ell}(t) = \frac{1}{t}\sum E\,\ell(t) = \frac{1}{t}\sum E[L + v(t)] = L$ 无偏估计

(设 $Ev^2(t) = \sigma^2$) $e(t) = \hat{l}_t - l$, $Ee(t) = 0$, $Ee^2(t) = E[(\frac{1}{t}\sum_{i=1}^{t} l(t) - l)^2]$

$= E[(\frac{1}{t}\sum_{i=1}^{t} l + v(t) - l)^2] = E[\frac{1}{t}\sum_{i=1}^{t} v(t)]^2$

$= E[\frac{1}{t^2}\cdot t\cdot \sigma^2] = \frac{1}{t}\sigma^2$ 方差

$J(\hat{l}(t)) = \sum_{i=1}^{t} (\underbrace{\frac{1}{t}\sum_{j=1}^{t} l(j) - l(i)}_{w(t)})^2$ ，其中 $w(t) = \frac{1}{t}\sum_{j=1}^{t} l + v(j) - l(i)$

$= l + \frac{1}{t}\sum_{j=1}^{t} v(j) - (l + v(i))$

$= \frac{1}{t}\sum_{j=1}^{t} v(j) - v(i)$

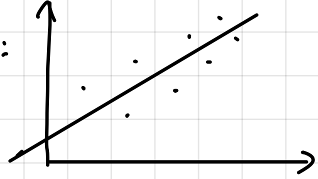
$\therefore J(\hat{l}(t)) = \sum_{i=1}^{t} w(t)^2 = \sum_{i=1}^{t} [\frac{1}{t}\sum_{j=1}^{t} v(j) - v(i)]^2$

$= \sum_{i=1}^{t} [\frac{1}{t}\sum_{j=1}^{t} v(j)]^2 - 2\sum_{i=1}^{t} [\frac{1}{t}\sum_{j=1}^{t} v(j)] v(i) + \sum_{i=1}^{t} v^2(i)$

$\therefore EJ(\hat{l}(t)) = t\cdot\frac{1}{t^2}\cdot t\cdot\sigma^2 - \frac{2}{t}\cdot t\sigma^2 + t\cdot\sigma^2 = (t-1)\sigma^2$

$\Longleftrightarrow E[\frac{1}{t-1} J(\hat{l}(t))] = \sigma^2$ ⟶ 可由此对白噪声进行估计

e.g 直线拟合:



$y=kx+b \qquad J=\sum_{i=1}^{t}(y(i)-(kx(i)+b))^2$

有 $\begin{cases} \frac{\partial J}{\partial k}=-\sum_{i=1}^{t}2(y(i)-(kx(i)+b))x(i)=0 \\ \frac{\partial J}{\partial b}=-\sum_{i=1}^{t}2(y(i)-(kx(i)+b))=0 \end{cases}$

$$\Rightarrow \begin{cases} \sum_{i=1}^{t}x^2(i)k+\sum_{i=1}^{t}x(i)b=\sum_{i=1}^{t}y(i)x(i) \\ \sum_{i=1}^{t}x(i)k+\sum_{i=1}^{t}b=\sum_{i=1}^{t}y(i) \end{cases} \xrightarrow[\text{形式}]{\text{矩阵}} \begin{bmatrix} \sum_{i=1}^{t}x^2(i) & \sum_{i=1}^{t}x(i) \\ \sum_{i=1}^{t}x(i) & \sum_{i=1}^{t}1 \end{bmatrix}\begin{bmatrix} k \\ b \end{bmatrix}=\begin{bmatrix} \sum_{i=1}^{t}y(i)x(i) \\ \sum_{i=1}^{t}y(i) \end{bmatrix}$$

$$\xrightarrow{\text{简化}} \sum_{i=1}^{t}\begin{bmatrix} x^2(i) & x(i) \\ x(i) & 1 \end{bmatrix}\begin{bmatrix} k \\ b \end{bmatrix}=\sum_{i=1}^{t}\begin{bmatrix} x(i) \\ 1 \end{bmatrix}\begin{bmatrix} x(i) & 1 \end{bmatrix}\begin{bmatrix} k \\ b \end{bmatrix}=\sum_{i=1}^{t}\begin{bmatrix} x(i) \\ 1 \end{bmatrix}y(i)$$

令 $\bar{x}(i)=\begin{bmatrix} x(i) \\ 1 \end{bmatrix}$, $\theta=\begin{bmatrix} k \\ b \end{bmatrix}$

$$\Rightarrow \underline{\left[\sum_{i=1}^{t}\bar{x}(i)\bar{x}^T(i)\right]}\theta=\sum_{i=1}^{t}\bar{x}(i)y(i)$$

$\hookrightarrow \underbrace{\begin{bmatrix} \bar{x}(1) & \bar{x}(2) & \cdots & \bar{x}(t) \end{bmatrix}}_{\in R^{2\times t}}\begin{bmatrix} \bar{x}^T(1) \\ \vdots \\ \bar{x}^T(t) \end{bmatrix}$

，令 $H_t=\begin{bmatrix} \bar{x}^T(1) \\ \vdots \\ \bar{x}^T(t) \end{bmatrix} \in R^{t\times 2} \Rightarrow H_t^T H_t\theta=H_t^T y(i)$

$$\Downarrow$$

$$\theta=\boxed{(H_t^T H_t)^{-1}H_t^T}\,y(i)$$

摩尔/加号逆 H_t^+

↳超定方程

从测量角度看:

$$\left.\begin{aligned} y(1)&=kx(1)+b+v(1)=\bar{x}^T(1)\theta+v(1) \\ y(2)&=kx(2)+b+v(2)=\bar{x}^T(2)\theta+v(2) \\ &\vdots \\ y(t)&=kx(t)+b+v(t)=\bar{x}^T(t)\theta+v(t) \end{aligned}\right\rangle \quad Y_t=H_t^T\theta$$

↳最小2乘解 $\|Y-H\theta\|_{min}$

则 $\theta=(H_t^T H_t)^{-1}H_t^T Y_t$

矩阵微分(矩阵求导)

$f(x)\in R$ $\qquad \frac{\partial f}{\partial x}=\begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}\in R^n$

$x\in R^n$

① $f(x)=\alpha^T x \qquad x=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \therefore f(x)=\alpha_1x_1+\alpha_2x_2+\dots+\alpha_nx_n \Rightarrow \frac{\partial f}{\partial x}=\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}=\alpha$

② $f(x)=x^TAx \qquad A=[a_{ij}]_{n\times n}=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{bmatrix}$，则 $x^TAx=\sum_i\sum_j a_{ij}x_ix_j=$

$x_1a_{11}x_1+x_1a_{12}x_2+\dots+x_1a_{1n}x_n+$
$x_2a_{21}x_1+x_2a_{22}x_2+\dots+x_2a_{2n}x_n+$
$\vdots$
$x_na_{n1}x_1+x_na_{n2}x_2+\dots+x_na_{nn}x_n$

$$\therefore \frac{\partial f}{\partial x}=\begin{bmatrix} a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n+a_{11}x_1+a_{21}x_2+\dots+a_{n1}x_n \\ a_{21}x_1+a_{22}x_2+\dots+a_{2n}x_n+a_{12}x_1+a_{22}x_2+\dots+a_{2n}x_n \\ \vdots \\ a_{n1}x_1+a_{n2}x_2+\dots+a_{nn}x_n+a_{1n}x_1+a_{2n}x_2+\dots+a_{nn}x_n \end{bmatrix}$$

$$=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & & & \vdots \\ \vdots & & & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}+\begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & & & \vdots \\ \vdots & & & \vdots \\ a_{1n} & \dots & \dots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=(A+A^T)x$$

If A is symmetric, $\frac{\partial f}{\partial x}=2Ax$，此时 A 正定. 有 Lyapunov function

$f(x)\in R^m$, $x\in R$ $\qquad f(x)=\begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{bmatrix}$，则 $\frac{\partial f}{\partial x}=\begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_2}{\partial x} & \dots & \frac{\partial f_n}{\partial x} \end{bmatrix}$

3、 $f(x)=\begin{bmatrix} \beta_1x \\ \beta_2x \\ \vdots \\ \beta_nx \end{bmatrix}=\beta x$. denote $\beta=\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}$，$\frac{\partial f}{\partial x}=\begin{bmatrix} \frac{\partial \beta_1x}{\partial x} & \frac{\partial \beta_2x}{\partial x} & \dots & \frac{\partial \beta_mx}{\partial x} \end{bmatrix}=[\beta_1\ \beta_2 \dots \beta_m]$

$=\beta^T$

相同形式!

3、 $x\in R^n$、 $f(x)\in R^m$

$$x=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},\quad f(x)=\begin{bmatrix}f_1(x)\\f_2(x)\\\vdots\\f_m(x)\end{bmatrix},\quad \frac{\partial f}{\partial x}=\begin{bmatrix}\frac{\partial f_1}{\partial x_1}&\frac{\partial f_2}{\partial x_1}&\cdots&\frac{\partial f_m}{\partial x_1}\\\frac{\partial f_1}{\partial x_2}&\frac{\partial f_2}{\partial x_2}&\cdots&\frac{\partial f_m}{\partial x_2}\\\vdots&&&\vdots\\\frac{\partial f}{\partial x_n}&\cdots&&\frac{\partial f_m}{\partial x_n}\end{bmatrix}\in R^{n\times m}$$

e.g. $\dot{x}=Ax$, $V(x(t))=x^TPx$, $\dot{V}=\dot{x}^TPx+x^T\dot{P}x$

推导:(链式法则推广) $x\in R^n$, $y\in R^p$, $z\in R^m$, $y=g(x)\in R^p$,

$z=h(y)=h(g(x))\in R^m$,

则 $\frac{\partial z}{\partial x}=\overset{n\times p}{\frac{\partial y}{\partial x}}\cdot\overset{p\times m}{\frac{\partial z}{\partial y}}\in R^{n\times m}$ (相容维数)

$f(x)=g(h(x))$, $\frac{\partial f}{\partial x}=\frac{\partial g_1}{\partial x}\cdot\frac{\partial f}{\partial g_1}+\frac{\partial g_2}{\partial x}\cdot\frac{\partial f}{\partial g_2}+\dots+\frac{\partial g_p}{\partial x}\cdot\frac{\partial f}{\partial g_p}=\sum_{i=1}^{p}\frac{\partial g_i}{\partial x}\cdot\frac{\partial f}{\partial g_i}$

$$\frac{\partial z}{\partial x}=\begin{bmatrix}\frac{\partial z_1}{\partial x_1}&\frac{\partial z_2}{\partial x_1}&\cdots&\frac{\partial z_m}{\partial x_1}\\\frac{\partial z_1}{\partial x_2}&\cdots&&\frac{\partial z_m}{\partial x_2}\\\vdots&&&\vdots\\\frac{\partial z_1}{\partial x_n}&\cdots&&\frac{\partial z_m}{\partial x_n}\end{bmatrix}=\begin{bmatrix}\sum_{i=1}^{p}\frac{\partial y_i}{\partial x_1}\cdot\frac{\partial z_1}{\partial y_i}&\sum_{i=1}^{p}\frac{\partial y_i}{\partial x_1}\cdot\frac{\partial z_2}{\partial y_i}&\cdots&\sum_{i=1}^{p}\frac{\partial y_i}{\partial x_1}\cdot\frac{\partial z_m}{\partial y_i}\\\sum_{i=1}^{p}\frac{\partial y_i}{\partial x_2}\cdot\frac{\partial z_1}{\partial y_i}&\sum_{i=1}^{p}\frac{\partial y_i}{\partial x_2}\cdot\frac{\partial z_2}{\partial y_i}&\cdots&\sum_{i=1}^{p}\frac{\partial y_i}{\partial x_2}\cdot\frac{\partial z_m}{\partial y_i}\\\vdots&&&\vdots\\\sum_{i=1}^{p}\frac{\partial y_i}{\partial x_n}\cdot\frac{\partial z_1}{\partial y_i}&\sum_{i=1}^{p}\frac{\partial y_i}{\partial x_n}\cdot\frac{\partial z_2}{\partial y_i}&\cdots&\sum_{i=1}^{p}\frac{\partial y_i}{\partial x_n}\cdot\frac{\partial z_m}{\partial y_i}\end{bmatrix}$$

$$=\begin{bmatrix}\frac{\partial y_1}{\partial x_1}&\frac{\partial y_2}{\partial x_1}&\cdots&\frac{\partial y_p}{\partial x_1}\\\frac{\partial y_1}{\partial x_2}&\frac{\partial y_2}{\partial x_2}&\cdots&\frac{\partial y_p}{\partial x_2}\\\vdots&\vdots&&\vdots\\\frac{\partial y_1}{\partial x_n}&\frac{\partial y_2}{\partial x_n}&\cdots&\frac{\partial y_p}{\partial x_n}\end{bmatrix}\begin{bmatrix}\frac{\partial z_1}{\partial y_1}&\frac{\partial z_2}{\partial y_1}&\cdots&\frac{\partial z_m}{\partial y_1}\\\frac{\partial z_1}{\partial y_2}&\frac{\partial z_2}{\partial y_2}&&\frac{\partial z_m}{\partial y_2}\\\vdots&\vdots&&\vdots\\\frac{\partial z_1}{\partial y_p}&\frac{\partial z_2}{\partial y_p}&\cdots&\frac{\partial z_m}{\partial y_p}\end{bmatrix}=\frac{\overset{R^p}{\partial y}}{\underset{R^n}{\partial x}}\cdot\frac{\overset{R^m}{\partial z}}{\partial y}$$

,证毕

(用向量的方式来思考)

Lyapwnov : $\begin{cases}\dot{x}=Ax\\V=x^TPx\end{cases}$ $\frac{\partial V}{\partial t}=\frac{\partial x}{\partial t}\cdot\frac{\partial V}{\partial x}=\dot{x}^T(P+P^T)x=x^TA^T(P+P^T)x$

$=x^T(A^TP+A^TP^T)x$

$=x^TA^TPx+\boxed{x^TPAx}$

$=x^T(A^TP+PA)x$

$x\in R^{m\times 1}$ ↓ 标量转置仍是标量

$$\dot{x}=\begin{bmatrix}\frac{dx_1}{dt}\\\frac{dx_2}{dt}\\\vdots\\\frac{dx_n}{dt}\end{bmatrix},\quad\boxed{\frac{\partial x}{\partial t}}=\begin{bmatrix}\frac{\partial x_1}{\partial t}&\frac{\partial x_2}{\partial t}&\cdots&\frac{\partial x_n}{\partial t}\end{bmatrix}=\boxed{\dot{x}^T}$$

当 $A^TP+PA<0$,稳定

$\Rightarrow A^TP+PA=-Q$,稳定

回推之前公式

由维数猜测

e.g. $f(x)=\underset{m\times n}{A}\underset{n\times 1}{x}\in R^{m\times 1}$, $x\in R^n$ $\therefore\frac{\partial f}{\partial x}\in R^{n\times m}\longrightarrow\frac{\partial f}{\partial x}=A^T$

证明: set $A=\begin{bmatrix}a_1\\a_2\\\vdots\\a_m\end{bmatrix}$, $a_i\in R^{1\times n}$ $\therefore f(x)=Ax=\begin{bmatrix}a_1x\\a_2x\\\vdots\\a_mx\end{bmatrix}$

(向量对标量求导)

$$\therefore\frac{\partial f}{\partial x}=\begin{bmatrix}\frac{\partial a_1x}{\partial x}&\frac{\partial a_2x}{\partial x}&\cdots&\frac{\partial a_mx}{\partial x}\end{bmatrix}=\begin{bmatrix}a_1^T&a_2^T&\cdots&a_m^T\end{bmatrix}=\begin{bmatrix}a_1\\a_2\\\vdots\\a_m\end{bmatrix}^T=A^T$$

,得证

系统辨识

$$y(1) = \theta_1 x_1(1) + \theta_2 x_2(1) + \dots + \theta_n x_n(1) + V(1)$$
$$\vdots \qquad \vdots \qquad \vdots$$
$$y(t) = \theta_1 x_1(t) + \theta_2 x_2(t) + \dots + \theta_n x_n(t) + V(t)$$

$$, J(t) = \sum_i^t (y(i) - \theta_1 x_1(i) - \dots - \theta_n x_n(i))^2$$

$\Longrightarrow$ 向量形式表达：$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}$; $Y_t = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(t) \end{bmatrix}$, $H_t = \begin{bmatrix} x^T(1) \\ x^T(2) \\ \vdots \\ x^T(t) \end{bmatrix} \in R^{t\times n}$

$\Rightarrow$ 简化：

$$J(t) = \sum_i^t (y(i) - x^T(i)\cdot\theta)^2 = (\qquad)^T \begin{bmatrix} y(1) - x^T(1)\theta \\ y(2) - x^T(2)\theta \\ \vdots \\ y(t) - x^T(t)\theta \end{bmatrix} = (Y_t - H_t\theta)^T (Y_t - H_t\theta)$$

(运用链式法则)

$$\frac{\partial J}{\partial \theta} = \frac{\partial (Y_t - H_t\theta)}{\partial \theta} \cdot \frac{\partial J}{\partial (Y_t - H_t\theta)} = -H_t^T \cdot 2(Y_t - H_t\theta) = 0 \Longleftrightarrow -H_t^T Y_t + H_t^T H_t \theta = 0$$
$$\Longleftrightarrow \theta = (H_t^T H_t)^{-1} H_t^T Y_t$$

$H_t^T Y = H_t^T H_t \theta \Longrightarrow \theta = (H_t^T H_t)^{-1} H_t^T Y$　　求逆计算量大，需要分解

正交化方法.

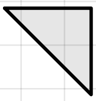
□QR：　$\underset{R^{m\times n}}{A} = \underset{R^{m\times m}}{Q} \; \underset{R^{n\times n}}{R}$　（其中Q是正交矩阵，R是上三角矩阵）

$$t\Big[\overset{n}{H_t} \;\; \overset{1}{Y_t}\Big] = QR = t\Big[\overset{t}{\qquad}\Big] \begin{bmatrix} \overbrace{R_1}^{n} & \overbrace{R_2}^{1} \\ 0 & R_3 \\ \multicolumn{2}{c}{0} \end{bmatrix} \begin{matrix} \}n \\ \}1 \\ \end{matrix} \xrightarrow{\because Q^TQ = I} Q^T[H_t \; Y_t] = R$$

考虑 $J = (Y - H_t\theta)^T (Y - H_t\theta) = (Y - H_t\theta)^T Q Q^T (Y - H_t\theta) = \| Q^T (Y - H_t\theta) \|^2$

$$= \left\| \begin{bmatrix} R_2 \\ R_3 \\ 0 \end{bmatrix} - \begin{bmatrix} R_1 \\ 0 \\ 0 \end{bmatrix} \theta \right\|^2 = \left\| \begin{bmatrix} R_2 - R_1\theta \\ R_3 \\ 0 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} R_2 - R_1\theta \\ R_3 \end{bmatrix} \right\|^2 = \|R_2 - R_1\theta\|^2 + R_3^2$$

易知 $R_2 - R_1\theta = 0$ 时，$\min J = R_3^2$

$R_1\theta = R_2 \Longrightarrow$  $[\quad] = [\quad] \Rightarrow$ 递推

要点：1、Q　　2. $R \to J_{min}$　　3. The dimension of the equation with respect to the parameter θ is reduced

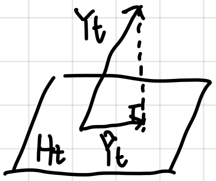
对 Y_t 进行估计，$Y_t = H_t\theta + V(t) \Rightarrow \hat{Y}(t) = H_t\hat{\theta}$

$$\hat{Y}_t = \underbrace{H_t (H_t^T H_t)^{-1} H_t^T}_{P} Y_t$$

P，有性质 $P^2 = H_t(H_t^T H_t)^{-1} H_t^T H_t (H_t^T H_t^{-1}) H_t^T = P$

同理 $P^n = P$（幂等矩阵 / 投影矩阵）

投影矩阵
几何解释：



□ 施密特正交化：思想：对上面的 $\hat{Y}_t = H_t(H_t^T H_t)^{-1}H_t^T Y_t$ 中 H_t 进行正交化

$$H_t = \begin{bmatrix} x^T(1)\\ x^T(2)\\ \vdots\\ x^T(t)\end{bmatrix} = \begin{array}{c} \begin{bmatrix} x_1(1) & x_2(1) & \cdots & x_n(1)\\ x_1(2) & x_2(2) & \cdots & x_n(2)\\ \vdots & & & \vdots\\ x_1(t) & x_2(t) & \cdots & x_n(t)\end{bmatrix} \\ \begin{array}{ccc} h_1 & h_2 & \qquad h_n \end{array}\end{array} = [h_1\ h_2\ \dots\ h_n]$$

$z_1 = h_1$

$z_2 = \lambda_{12} z_1 + h_2$

$\left(z_1 \perp z_2 \quad \therefore z_1 \cdot z_2 = h_1^T(h_1 + \lambda_{12} h_2) = \|h_1\|^2 + \lambda_{22} h_1^T h_2 = 0 \right.$

$\left. \therefore \lambda_{22} = -\frac{\|h_1\|^2}{h_1^T h_2} \right)$

$z_3 = \lambda_{13} z_1 + \lambda_{23} z_2 + h_3$

$$\Longrightarrow [h_1\ h_2\ \dots\ h_n] = [z_1\ z_2\ \dots\ z_n]\begin{bmatrix} 1 & -\lambda_{12} & -\lambda_{13} & \vdots\\ 0 & 1 & -\lambda_{23} & \vdots\\ \vdots & & 1 & \vdots\\ 0 & \cdots & \ddots & 1\end{bmatrix} = \left[\frac{z_1}{\|z_1\|}\ \frac{z_2}{\|z_2\|} \cdots \frac{z_n}{\|z_n\|}\right]\begin{bmatrix} \|z_1\| & & & \\ & \|z_2\| & & \\ & & \ddots & \\ & & & \|z_n\|\end{bmatrix}\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$$

$$\Longrightarrow \underset{t\times n}{H_t} = \underset{t\times n}{Z}\ \underset{n\times n}{R}$$

$$H_t\theta = Y_t \Rightarrow ZR\theta = Y_t \Rightarrow Z^T Z R\theta = Z^T Y_t$$

$$\Longrightarrow R\theta = Z^T Y_t \qquad \nabla[\ ] = [\ ]$$

最小二乘统计性质：

$y(t) = x^T(t)\theta + v(t) \Longrightarrow Y_t = H_t\theta + V(t)$，其中 $Ev(t) = 0$

（白噪声假设）

$\hat{\theta}(t) = (H_t^T H_t)^{-1} H_t^T Y_t \rightarrow \hat{\theta}(t) = (H_t^T H_t)^{-1} H_t^T(H_t\theta + v(t)) = \theta + (H_t^T H_t)^{-1} H_t^T v(t)$

① $E\hat{\theta}(t) = \theta$

$$\hat{\theta}(t) = (H_t^T H_t)^{-1} H_t^T(H_t\theta + V(t)) = (H_t^T H_t)^{-1} H_t^T H_t\theta + (H_t^T H_t)^{-1} H_t^T v(t)$$

（推导 $EAX = AEX$，设 $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & & & a_{2n}\\ \vdots & & & \vdots\\ a_{m1} & \dots & \dots & a_{mn}\end{bmatrix}$，

（已知 $E\begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix} = \begin{bmatrix} Ex_1\\ Ex_2\\ \vdots\\ Ex_n\end{bmatrix}$）

则 $EAX = E\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n\\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n\\ \vdots\\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n\end{bmatrix} \overset{\text{定义}}{\underset{\text{期望性质}}{=\!=\!=}} \begin{bmatrix} a_{11}Ex_1 + a_{12}Ex_2 + \dots + a_{1n}Ex_n\\ a_{21}Ex_1 + a_{22}Ex_2 + \dots + a_{2n}Ex_n\\ \vdots \qquad\qquad \vdots\\ a_{m1}Ex_1 + a_{m2}Ex_2 + \dots + a_{mn}Ex_n\end{bmatrix} = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}\begin{bmatrix} Ex_1\\ Ex_2\\ \vdots\\ Ex_n\end{bmatrix}$

$= AEX$ ）

$\therefore E\hat{\theta}(t) = (H_t^T H_t)^{-1} H_t^T H_t E\theta + (H_t^T H_t)^{-1} H_t^T Ev(t) = E\theta = \theta$，无偏估计

② $e(t) = \hat{\theta}(t) - \theta$，$Ee(t) = 0$，$Cov[e(t)]$ （Covariance：EXX^T）

$e(t) = \hat{\theta}(t) - \theta = (H_t^T H_t)^{-1} H_t^T v(t)$ （$= (H_t^T H_t)^{-1} H_t^T Y_t - \theta = (H_t^T H_t)^{-1} H_t^T H_t(\theta + v(t)) - \theta$

$$E[e(t)\,e(t)^T] = E[\underbrace{(H_t^T H_t)^{-1} H_t^T}_{A_1} v(t)v(t)^T \underbrace{H_t(H_t^T H_t)^{-1}}_{A_2}] = A_1 E[v(t)v(t)^T] A_2$$

$$= \sigma^2 A_1 A_2$$

$$= \sigma^2 (H_t^T H_t)^{-1}$$

我们希望对数据有所选择

③ $J(\theta)=\sum_{i=1}^{t}(y(i)-x^T(i)\theta)^2=\|Y_t-H_t\theta\|^2$

$$\Rightarrow J(\hat{\theta})=\|Y_t-H_t\hat{\theta}(t)\|^2=\|Y_t-H_t\cdot(H_t^TH_t)^{-1}H_t^TY_t\|^2=\|(I-H_t(H_t^TH_t)^{-1}H_t^T)(H_t\theta+V_t)\|_2$$

$$=\|(I-H_t(H_t^TH_t)^{-1}H_t^T)V_t\|^2=\|(I-Q)V_t\|^2$$

$$\therefore E(J(\hat{\theta}))=E(\|(I-Q)V_t\|^2)=E(V_t^T(I-Q)^T(I-Q)V_t)=E[V_t^T(I-Q-Q^T+Q^TQ)V_t]$$

（其中 $Q=H_t(H_t^TH_t)^{-1}H_t^T$，$Q=Q^T$（对称阵），$Q^2=Q$；且有性质 $tr(BA)=tr(AB)$）

$$Q^2=Q^TQ=H_t(H_t^TH_t)^{-1}H_t^TH_t(H_t^TH_t^{-1})H_t^T=Q$$

$\therefore$ 上式 $=E[V_t^T(I-Q)V_t]=E[tr(V_t^T(I-Q)V_t)]=E[tr(V_t^TV_t(I-Q))]$

$\downarrow$
标量

（取迹和求期望可交换顺序：e.g. $E\,tr(A)=E(a_{11}+a_{22}+\dots+a_{nn})=Ea_{11}+Ea_{22}+\dots+Ea_{nn}$）

$$tr\,E(A)=tr\begin{bmatrix}Ea_{11}&Ea_{12}&\dots&Ea_{1n}\\Ea_{21}&&&\vdots\\\vdots&&&\vdots\\Ea_{n1}&Ea_{n2}&\dots&Ea_{nn}\end{bmatrix}=Ea_{11}+Ea_{22}+\dots+Ea_{nn}$$

$$\therefore E[tr(V_t^TV_t(I-Q))]=tr[E(V_t^TV_t(I-Q)]=tr[E(V_t^TV_t)(I-Q)]$$

$$=tr[\sigma^2(I-Q)]=\sigma^2\,tr(I-Q)=\sigma^2(tr\,I-tr\,Q)=\sigma^2(t-n)$$

$(tr\,Q=tr(H_t(H_t^TH_t)^{-1}H_t^T)=tr[(H_t^TH_t^{-1})H_t^TH_t]=tr\,I\in R^n=n)$

$$\Rightarrow E\frac{J(\hat{\theta})}{t-n}=\sigma^2$$

Recursive least Square Identification

系统误差模型: $A(z^{-1})y(t)=B(z^{-1})u(t)+v(t)$ $\begin{cases} A(z^{-1})=1+a_1z^{-1}+\dots+a_{n_a}z^{-n_a} \\ B(z^{-1})=b_1z^{-1}+\dots+b_{n_b}z^{-n_b} \end{cases}$

（与当前时刻输入无关）

↓ 即考虑数据: $y(1), y(2), \dots, y(t)$, $u(1), u(2), \dots, u(t-1)$

⇒ 差分形式 $y(t)+a_1y(t-1)+\dots+a_{n_a}y(t-n_a)=b_1u(t-1)+b_2u(t-2)+\dots+b_{n_b}u(t-n_b)+v(t)$

有 $y(t)=-a_1y(t-1)-a_2y(t-2)-a_{n_a}y(t-n_a)+b_1u(t-1)+b_2u(t-2)+\dots+b_{n_b}(t-n_b)+v(t)$

令 $\theta=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_a} \\ b_1 \\ \vdots \\ b_{n_b} \end{bmatrix}$, $\varphi(t)=\begin{bmatrix} -y(t-1) \\ -y(t-2) \\ \vdots \\ -y(t-n_a) \\ u(t-1) \\ \vdots \\ u(t-n_b) \end{bmatrix}$ $\Longrightarrow$ $y(t)=\varphi^T(t)\theta+v(t)$

identification model

指标函数: $J=\sum_{i=1}^{t}(y(i)-\varphi^T(i)\theta)^2$

$=(Y_t-H_t\theta)^T(Y_t-H_t\theta)$

向量 → $Y_t=\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(t) \end{bmatrix}$, $H_t=\begin{bmatrix} \varphi^T(1) \\ \varphi^T(2) \\ \vdots \\ \varphi^T(t) \end{bmatrix}$ （对小于0的数据: 1、取零或任意值 2、运行一段时间后再辨识

$\Longrightarrow \hat{\theta}(t)=(H_t^TH_t)^{-1}H_t^TY_t \longrightarrow P(t)$: $P(t)=(H_t^TH_t)^{-1}$,且易知正定

递推形式: $\hat{\theta}(t)=(H_t^TH_t)^{-1}H_t^TY_t$

$\hat{\theta}(t-1)=(H_{t-1}^TH_{t-1})^{-1}H_{t-1}^TY_{t-1}$

分析得 $H_t=\begin{bmatrix} H_{t-1} \\ \varphi^T(t) \end{bmatrix}$, $Y_t=\begin{bmatrix} Y_{t-1} \\ y(t) \end{bmatrix}$

代入得 $\hat{\theta}(t)=P(t)[H_{t-1}^T \ \ \varphi(t)]\begin{bmatrix} Y_{t-1} \\ y(t) \end{bmatrix}=P(t)(H_{t-1}^TY_{t-1}+\varphi(t)y(t))$

$\overset{\text{还原}}{\underset{\text{形式}}{=\!=}} P(t)\cdot(P^{-1}(t-1)\cdot P(t-1)H_{t-1}^TY_{t-1}+\varphi(t)y(t))$

$=P(t)P^{-1}(t-1)\hat{\theta}(t-1)+P(t)\varphi(t)y(t)$ 之后简化$P(t)$, $P(t-1)$

$\Big(P^{-1}(t)=H_t^TH_t=[H_{t-1}^T \ \ \varphi(t)]\begin{bmatrix} H_{t-1} \\ \varphi^T(t) \end{bmatrix}=H_{t-1}^TH_{t-1}+\varphi(t)\varphi^T(t)=P^{-1}(t-1)+\varphi(t)\varphi^T(t)$

即 $P^{-1}(t)=P^{-1}(t-1)+\varphi(t)\varphi^T(t)$ $\Big)$

∴有 $I=P(t)P^{-1}(t-1)+P(t)\varphi(t)\varphi^T(t) \Longrightarrow P(t)P^{-1}(t-1)=I-P(t)\varphi(t)\varphi^T(t)$

∴原式 $=(I-P(t)\varphi(t)\varphi^T(t))\hat{\theta}(t-1)+P(t)\varphi(t)y(t)$

$=\hat{\theta}(t-1)+P(t)\varphi(t)(y(t)-\varphi^T(t)\hat{\theta}(t-1))$

$$\Longrightarrow \begin{cases} \hat{\theta}(t)=\hat{\theta}(t-1)+\overset{\text{gain}}{\underline{P(t)\varphi(t)}}(\overset{\text{measurement}}{\underline{y(t)}}-\underline{\varphi^T(t)\hat{\theta}(t-1)}) \\ P^{-1}(t)=P^{-1}(t-1)+\varphi(t)\varphi^T(t) \end{cases}$$

↳上一步估计值对当前步的预测

解释: gain : $P(t)\varphi(t)$

innovation: $y(t)-\varphi^T(t)\hat{\theta}(t-1)$

两条公式针对ARX模型可独立进行，其余模型必须交替进行

↓ next page: 初值选择, 性质

P(t)更新示例: $P^{-1}(1)=P^{-1}(0)+\varphi(1)\varphi^T(1)$

$$P^{-1}(2)=P^{-1}(0)+\varphi(2)\varphi^T(2)=P^{-1}(0)+\varphi(1)\varphi^T(1)+\varphi(2)\varphi^T(2)$$

$$\vdots$$

$$P^{-1}(t)=P^{-1}(0)+\varphi(1)\varphi^T(1)+\dots+\varphi(t)\varphi^T(t)$$

取 $P(0)=PI$，P 取 10^6 $\Rightarrow$ 初值较大

矩阵求逆: $(A+BC)^{-1}=A^{-1}-A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1}$ 进行简化

$$\Longrightarrow P(t)=(P^{-1}(t-1)+\varphi(t)\varphi^T(t))^{-1}$$

标量

$$=P(t-1)-P(t-1)\varphi(t)\left(I+\varphi^T(t)P(t-1)\varphi(t)\right)^{-1}\varphi^T(t)P(t-1)$$

$L(t)$

半正定矩阵

$$=P(t-1)-\frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{I+\varphi^T(t)P(t-1)\varphi(t)}$$

从形式可知，P(t)逐渐减小

$$P(t)\varphi(t)=P(t-1)\varphi(t)-\frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)\varphi(t)}{I+\varphi^T(t)P(t-1)\varphi(t)}$$

标量

$$=P(t-1)\varphi(t)\left(1-\frac{\varphi^T(t)P(t-1)\varphi(t)}{I+\varphi^T(t)P(t-1)\varphi(t)}\right)$$

$$=\frac{P(t-1)\varphi(t)}{I+\varphi^T(t)P(t-1)\varphi(t)}$$

$$\Longrightarrow\begin{cases}\hat{\theta}(t)=\hat{\theta}(t-1)+L(t)(y(t)-\varphi^T(t)\hat{\theta}(t-1))\\ L(t)=\dfrac{P(t-1)\varphi(t)}{I+\varphi^T(t)P(t-1)\varphi(t)}\\ P(t)=P(t-1)-L(t)\varphi^T(t)P(t-1)=[I-L(t)\varphi^T(t)]P(t-1)\end{cases}$$

$\boxed{\text{最常用RLS}}$

更新步骤:

$$\begin{tikzcd}
\hat{\theta}(0) \arrow[r] & \hat{\theta}(1) \arrow[r] & \hat{\theta}(2) \arrow[r] & \hat{\theta}(3) \\
P(0) \arrow[r] \arrow[dr] & L(1) \arrow[u] \arrow[d] & L(2) \arrow[u] \arrow[d] & L(3) \arrow[u] \\
 & P(1) \arrow[r] \arrow[ur] & P(2) \arrow[ur] &
\end{tikzcd}$$

考虑指标: $\begin{cases}e(t)=y(t)-\varphi^T(t)\hat{\theta}(t-1)\\ \varepsilon(t)=y(t)-\varphi^T(t)\hat{\theta}(t)\end{cases}$

$L(t)=P(t)\varphi(t)$

$$\varepsilon(t)=y(t)-\varphi^T(t)(\hat{\theta}(t-1)+L(t)e(t))$$

$$=y(t)-\varphi^T(t)\hat{\theta}(t-1)-\varphi^T(t)L(t)e(t)=[I-\varphi^T(t)L(t)]e(t)$$

$$=\left(1-\frac{\varphi^T(t)P(t-1)\varphi(t)}{I+\varphi^T(t)P(t-1)\varphi(t)}\right)e(t)=\frac{1}{I+\varphi^T(t)P(t-1)\varphi(t)}e(t)$$

$\varepsilon(t)$，$e(t)$ 同号；

且有 $e(t)=\dfrac{\varepsilon(t)}{I-\varphi^T(t)P(t)\varphi(t)}$ 可知 $0<\varphi^T(t)P(t)\varphi(t)<1$

将两式相乘：$(I-\varphi^T(t)P(t)\varphi(t))(I+\varphi^T(t)P(t-1)\varphi(t))=1$

$$\Rightarrow I-\varphi^T(t)P(t)\varphi(t)=\frac{1}{I+\varphi^T(t)P(t-1)\varphi(t)}$$

新的迭代算法

（上述的 ARX 模型递推最小二乘，仍然存在数据饱和问题）

$P(t)=P(t-1)-L(t)\varphi^T(t)P(t-1)=[I-L(t)\varphi^T(t)]P(t-1)$，不断减小

则 $\hat\theta(t)=\hat\theta(t-1)+L(t)(y(t)-\varphi^T(t)\hat\theta(t-1))$，gain 将不起作用

出现 data saturation 的原因：

$\hat\theta(t)=(H_t^TH_t)^{-1}H_t^TY_t$，$P^{-1}(t)=H_t^TH_t$，有 $\hat\theta(t)=P(t)H_t^TY_t$

$H_t=\begin{bmatrix}H_{t-1}\\ \varphi^T(t)\end{bmatrix}$，$Y_t=\begin{bmatrix}Y_{t-1}\\ y(t)\end{bmatrix}$，新数据和先前数据权重相同，

改进思路：削弱原先数据影响：$H_t=\begin{bmatrix}\rho H_{t-1}\\ \varphi^T(t)\end{bmatrix}$，$Y_t=\begin{bmatrix}\rho Y_{t-1}\\ y(t)\end{bmatrix}$，$\rho\in(0,1)$

$$\therefore \hat\theta(t)=P(t)H_t^TY_t=P(t)[\rho H_{t-1}^T\ \ \varphi(t)]\begin{bmatrix}\rho Y_{t-1}\\ y(t)\end{bmatrix}=P(t)(\rho^2H_{t-1}^TY_{t-1}+\varphi(t)y(t))$$

$$=P(t)\rho^2H_{t-1}^TY_{t-1}+P(t)\varphi(t)y(t)$$

与 $\hat\theta(t-1)=P(t-1)H_{t-1}^TY_{t-1}$ 相关联：

$$\Rightarrow \hat\theta(t)=P(t)\rho^2P^{-1}(t-1)\hat\theta(t-1)+P(t)\varphi(t)y(t)$$

其中 $P^{-1}(t)=H_t^TH_t=[\rho H_{t-1}^T\ \ \varphi(t)]\begin{bmatrix}\rho H_{t-1}\\ \varphi^T(t)\end{bmatrix}=\rho^2P^{-1}(t-1)+\varphi(t)\varphi^T(t)$

$I=\rho^2P(t)P^{-1}(t-1)+P(t)\varphi(t)\varphi^T(t)$　　　记 $\rho^2=\lambda$

$$\therefore\begin{cases}\hat\theta(t)=[I-P(t)\varphi(t)\varphi^T(t)]\hat\theta(t-1)+P(t)\varphi(t)y(t)\\ \quad\quad=\hat\theta(t-1)+P(t)\varphi(t)[y(t)-\varphi^T(t)\hat\theta(t-1)]\\ P^{-1}(t)=\lambda P^{-1}(t-1)+\varphi(t)\varphi^T(t)\end{cases}$$

形式相同

实际上是对代价函数的修改：

$$J=\sum_{i=1}^{t}(y(i)-\varphi^T(i)\theta)^2\Longrightarrow J=\sum_{i=1}^{t}\overset{(\rho^{2i-2})}{\lambda^{i-1}}(y(i)-\varphi^T(i)\theta)^2$$

$$H_t=\begin{bmatrix}\varphi^T(1)\\ \varphi^T(2)\\ \vdots\\ \varphi^T(t)\end{bmatrix},\ Y_t=\begin{bmatrix}y(1)\\ y(2)\\ \vdots\\ y(t)\end{bmatrix}\Longrightarrow H_t'=\begin{bmatrix}\rho^{t-1}\varphi^T(1)\\ \rho^{t-2}\varphi^T(2)\\ \vdots\\ \varphi^T(t)\end{bmatrix},\ Y_t'=\begin{bmatrix}\rho^{t-1}y(1)\\ \rho^{t-2}y(2)\\ \vdots\\ y(t)\end{bmatrix}$$

$$\hookrightarrow H_t'=\begin{bmatrix}\rho^{t-1}&&&\\ &\rho^{t-2}&&\\ &&\ddots\ \rho^1&\\ &&&\rho^0\end{bmatrix}H_t$$

$\Lambda^{\frac{1}{2}}$

$Y_t'=\Lambda^{\frac{1}{2}}Y_t$

$$\Lambda=\begin{bmatrix}\lambda^{t-1}&&&\\ &\lambda^{t-2}&&\\ &&\ddots\ \lambda^1&\\ &&&1\end{bmatrix}$$

J 向量化：$J=(Y_t'-H_t'\theta)^T(Y_t'-H_t'\theta)$

$=(Y_t-H_t\theta)^T\Lambda_t(Y_t-H_t\theta)$

$\Rightarrow \theta=(H_t'^TH_t')^{-1}H_t'^TY_t'$

$=(H_t^T\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}H_t)^{-1}H_t^T\cdot\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}Y_t$

$=(H_t^T\Lambda H_t)^{-1}H_t^T\Lambda Y_t$

$$\left(\frac{\partial J}{\partial \hat{\theta}} = \frac{\partial(Y_t - H_t\hat{\theta})}{\partial\hat{\theta}} \cdot \frac{\partial J}{\partial(Y_t - H_t\hat{\theta})} = -H_t^T \cdot \left(2\Lambda_t (Y_t - H_t\hat{\theta})\right) = 0 \right.$$

$$\Longleftrightarrow \quad H_t^T \Lambda_t Y_t = H_t^T \Lambda_t H_t \hat{\theta} \quad \Longrightarrow \quad \hat{\theta} = (H_t^T \Lambda_t H_t)^{-1} H_t \Lambda_t Y_t$$

$$\hat{\theta}(t) = \underbrace{\left[\sum_{i=1}^{t} \lambda^{t-i} \varphi(i) \varphi^T(i)\right]^{-1}}_{P(t)} \sum_{i=1}^{t} \lambda^{t-i} \varphi(i) y(i)$$

$$= P(t)\left[\sum_{i=1}^{t-1} \lambda^{t-i} \varphi(i) y(i) + \varphi(t) y(t)\right] = P(t) \cdot \left[\lambda P^{-1}(t-1) \sum_{i=1}^{t-1} \lambda^{t-i-1} P(t-1) \varphi(i) y(i) + \varphi(t) y(t)\right]$$

$$= \lambda P(t) P^{-1}(t-1)\, \hat{\theta}(t-1) + P(t) \varphi(t) y(t) \underset{\uparrow}{=} [1 - P(t)\varphi(t)\varphi^T(t)]\, \hat{\theta}(t-1) + P(t)\varphi(t) y(t)$$

$$P^{-1}(t) = \sum_{i=1}^{t} \lambda^{t-i} \varphi(i) \varphi^T(i) \Rightarrow \lambda \sum_{i=1}^{t-1} \lambda^{t-i-1} \varphi(i) \varphi^T(i) + \varphi(t)\varphi^T(t) = \lambda P^{-1}(t-1) + \varphi(t)\varphi^T(t)$$

$$\Longrightarrow \begin{cases} \hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\varphi(t)\left(y(t) - \varphi^T(t)\hat{\theta}(t-1)\right) \\ P^{-1}(t) = \lambda P^{-1}(t-1) + \varphi(t)\varphi^T(t) \end{cases}$$

$$\Downarrow \quad (A+BC)^{-1} = A^{-1} - A^{-1}B(I + CA^{-1}B)^{-1} CA^{-1}$$

$$\therefore P(t) = \left(\lambda P^{-1}(t-1) + \varphi(t)\varphi^T(t)\right)^{-1} = \frac{1}{\lambda}\left[P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{\lambda + \varphi^T(t)P(t-1)\varphi(t)}\right]$$

$$\Downarrow$$

$$L(t) = P(t)\varphi(t) = \frac{1}{\lambda}\left[P(t-1)\varphi(t) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)\varphi(t)}{\lambda + \varphi^T(t)P(t-1)\varphi(t)}\right]$$

$$= \frac{1}{\lambda} P(t-1)\varphi(t) \frac{\lambda}{\lambda + \varphi^T(t)P(t-1)\varphi(t)} = \frac{P(t-1)\varphi(t)}{\lambda + \varphi^T(t)P(t-1)\varphi(t)}$$

回代： $P(t) = \frac{1}{\lambda}\left[P(t-1) - L(t)\varphi^T(t)P(t-1)\right] = \frac{1}{\lambda}\left[I - L(t)\varphi^T(t)\right]P(t-1)$

$$\Longrightarrow \begin{cases} \hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\varphi(t)\left(y(t) - \varphi^T(t)\hat{\theta}(t-1)\right) \\ P(t) = \frac{1}{\lambda}\left[I - L(t)\varphi^T(t)\right]P(t-1) \\ L(t) = \frac{P(t-1)\varphi(t)}{\lambda + \varphi^T(t)P(t-1)\varphi(t)} \end{cases}$$

☆

$$\varepsilon(t) = y(t) - \varphi^T(t)\hat{\theta}(t) = y(t) - \varphi^T(t)\left(\hat{\theta}(t-1) + L(t)e(t)\right)$$

$$= [1 - \varphi^T(t)L(t)]e(t) = \frac{\lambda}{\lambda + \varphi^T(t)P(t-1)\varphi(t)} e(t)$$

$$\left\{ \begin{array}{l} \varepsilon(t) = \left(1 - \varphi^T(t)P(t)\varphi(t)\right)e(t) \\ e(t) = \left[1 + \frac{1}{\lambda} \cdot \varphi^T(t)P(t-1)\varphi(t)\right]\varepsilon(t) \end{array} \right\rangle \quad \varepsilon(t) \leq e(t)$$

↓

有限数据窗的递推最小二乘

$$J=\sum_{i=1}^{t}(y(i)-\varphi^T(i)\theta)^2 \Longrightarrow J=\sum_{i=t-p+1}^{t}(y(i)-\varphi^T(i)\theta)^2$$ （数据窗长度p）

$$=[\quad]^T\begin{bmatrix} y(t-p+1)-\varphi^T(t-p+1)\theta \\ y(t-p+2)-\varphi^T(t-p+2)\theta \\ \vdots \\ y(t)-\varphi^T(t)\theta \end{bmatrix}$$

$$=(Y_{t,p}-H_{t,p}\theta)^T(Y_{t,p}-H_{t,p}\theta)$$

$$\therefore \hat\theta(t)=[H_{t,p}^TH_{t,p}]^{-1}[H_{t,p}^TY_{t,p}] \quad ; \quad \hat\theta(t-1)=[H_{t-1,p}^TH_{t-1,p}]^{-1}[H_{t-1,p}^TY_{t-1,p}]$$

$$P(t)=(H_{t,p}^TH_{t,p})^{-1}, \quad P(t-1)=(H_{t-1,p}^TH_{t-1,p})^{-1}$$

$$H_{t,p}^TY_{t,p}=\sum_{i=t-p+1}^{t}\varphi(i)y(i) \quad H_{t-1,p}^TY_{t-1,p}=\sum_{i=t-p}^{t}\varphi(i)y(i)$$

$$\hat\theta(t)=P(t)\sum_{i=t-p+1}^{t}\varphi(i)y(i)=P(t)\left[\sum_{i=t-p}^{t-1}\varphi(i)y(i)+\varphi(t)y(t)-\varphi(t-p)y(t-p)\right]$$

$$=P(t)\cdot\left[P^{-1}(t-1)\;P(t-1)\sum_{i=t-p}^{t-1}\varphi(i)y(i)+\varphi(t)y(t)-\varphi(t-p)y(t-p)\right]$$

$$=P(t)[P^{-1}(t-1)\;\hat\theta(t-1)+\varphi(t)y(t)-\varphi(t-p)y(t-p)]$$

$$P^{-1}(t)=H_{t,p}^TH_{t,p}=\sum_{i=t-p+1}^{t}\varphi(i)\varphi^T(i) \qquad P^{-1}(t-1)=\sum_{i=t-p}^{t-1}\varphi(i)\varphi^T(i)$$

$$=\sum_{i=t-p}^{t-1}\varphi(i)\varphi^T(i)+\varphi(t)\varphi^T(t)-\varphi(t-p)\varphi^T(t-p)$$

$$\Longrightarrow P^{-1}(t)=P^{-1}(t-1)+\varphi(t)\varphi^T(t)-\varphi(t-p)\varphi^T(t-p)$$

$$\Longrightarrow I=P(t)P^{-1}(t-1)+P(t)\varphi(t)\varphi^T(t)-P(t)\varphi(t-p)\varphi^T(t-p)$$

$$\therefore \hat\theta(t)=[I-P(t)\varphi(t)\varphi^T(t)+P(t)\varphi(t-p)\varphi^T(t-p)]\hat\theta(t-1)+P(t)\varphi(t)y(t)-P(t)\varphi(t-p)y(t-p)$$

$$=\hat\theta(t-1)+P(t)\varphi(t)[y(t)-\varphi^T(t)\hat\theta(t-1)]-P(t)\varphi(t-p)[y(t-p)-\varphi^T(t-p)\hat\theta(t-1)]$$

简化P(t)表达: ① $P^{-1}(t)=P^{-1}(t-1)+[\varphi(t)\;\;\varphi(t-p)]\begin{bmatrix}\varphi^T(t)\\ \varphi^T(t-p)\end{bmatrix}$ $\quad ((A+BC)^{-1}=A^{-1}-A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1})$

(dim: n×n)

$$\therefore P(t)=P(t-1)-P(t-1)\cdot[\varphi(t)\;\;\varphi(t-p)]\left(I+\begin{bmatrix}\varphi^T(t)\\ \varphi^T(t-p)\end{bmatrix}P(t-1)[\varphi(t)\;\;\varphi(t-p)]\right)^{-1}\begin{bmatrix}\varphi^T(t)\\ \varphi^T(t-p)\end{bmatrix}P(t-1)$$

令 $Q^{-1}(t)=P^{-1}(t-1)+\varphi(t)\varphi^T(t)$

$P^{-1}(t)=Q^{-1}(t)+\varphi(t-p)\varphi^T(t-p) \quad ((A+BC)^{-1}=A^{-1}-A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1})$

② $\Rightarrow \; P(t)=Q(t)+\dfrac{Q(t)\varphi(t-p)\varphi^T(t-p)Q(t)}{1-\varphi^T(t-p)Q(t)\varphi(t-p)}, \quad Q(t)=P(t-1)-\dfrac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{1+\varphi^T(t)P(t-1)\varphi(t)}$

两次利用求逆引理

$$\Longrightarrow Y_{t-1,p-1}=\begin{bmatrix} y(t-p+1)\\ y(t-p+2)\\ \vdots \\ y(t-1)\end{bmatrix} \Rightarrow Y_{t,p}=\begin{bmatrix}Y_{t-1,p-1}\\ y(t)\end{bmatrix}, \quad Y_{t-1,p}=\begin{bmatrix}y(t-p)\\ Y_{t-1,p-1}\end{bmatrix}$$

$$H_{t-1,p-1}=\begin{bmatrix} \varphi^T(t-p+1)\\ \varphi^T(t-p+2)\\ \vdots \\ \varphi^T(t-1)\end{bmatrix} \Longrightarrow H_{t,p}=\begin{bmatrix}H_{t-1,p-1}\\ \varphi^T(t)\end{bmatrix}, \quad H_{t-1,p}=\begin{bmatrix}\varphi^T(t-p)\\ H_{t-1,p-1}\end{bmatrix}$$

$$P_\alpha(t-1)=(H^T_{t-1,p-1}H_{t-1,p-1})^{-1}, \quad \alpha(t-1)=P_\alpha(t-1)H^T_{t-1,p-1}Y_{t-1,p-1}$$

（可知 $\hat\theta(t)\frown P_\alpha(t-1)\frown\hat\theta(t-1)$）☆

$$\hat\theta(t)=P(t)[H^T_{t-1,p-1}\ \ \varphi(t)]\begin{bmatrix}Y_{t-1,p-1}\\ y(t)\end{bmatrix}=P(t)[P_\alpha^{-1}(t-1)P_\alpha(t-1)H^T_{t-1,p-1}Y_{t-1,p-1}+\varphi(t)y(t)]$$

$$=P(t)[P_\alpha^{-1}(t-1)\cdot\alpha(t-1)+\varphi(t)y(t)]$$

（有 $P^{-1}(t)=[H^T_{t-1,p-1}\ \ \varphi(t)]\begin{bmatrix}H_{t-1,p-1}\\ \varphi^T(t)\end{bmatrix}=H^T_{t-1,p-1}H_{t-1,p-1}+\varphi(t)\varphi^T(t)$
$=P_\alpha^{-1}(t-1)+\varphi(t)\varphi^T(t)$）

$$\Rightarrow P^{-1}(t)=P_\alpha^{-1}(t-1)+\varphi(t)\varphi^T(t)$$

$$\therefore I=P(t)P_\alpha^{-1}(t-1)+P(t)\varphi(t)\varphi^T(t)$$

$$\therefore \hat\theta(t)=[I-P(t)\varphi(t)\varphi^T(t)]\alpha(t-1)+P(t)\varphi(t)y(t)$$

$$=\alpha(t-1)+P(t)\varphi(t)(y(t)-\varphi^T(t)\alpha(t-1))$$

$$\hat\theta(t-1)=P(t-1)H^T_{t-1,p}Y_{t-1,p}=P(t-1)[\varphi(t-p)\ \ H^T_{t-1,p-1}]\begin{bmatrix}y(t-p)\\ Y_{t-1,p-1}\end{bmatrix}$$

$$=P(t-1)[P_\alpha^{-1}(t-1)P_\alpha(t-1)H^T_{t-1,p-1}Y_{t-1,p-1}+\varphi(t-p)y(t-p)]$$

$$=P(t-1)P_\alpha^{-1}(t-1)\alpha(t-1)+P(t-1)\varphi(t-p)y(t-p)$$

$\Downarrow$ 解 $\alpha(t-1)$

$$P^{-1}(t-1)\hat\theta(t-1)=P_\alpha^{-1}(t-1)\alpha(t-1)+\varphi(t-p)y(t-p)$$

$$P_\alpha(t-1)P^{-1}(t-1)\theta(t-1)=\alpha(t-1)+P_\alpha(t-1)\varphi(t-p)y(t-p)$$

$$\therefore \alpha(t-1)=P_\alpha(t-1)P^{-1}(t-1)\hat\theta(t-1)-P_\alpha(t-1)\varphi(t-p)y(t-p)$$

（求解 $P^{-1}(t-1)$：$P^{-1}(t-1)=[\varphi(t-p)\ \ H^T_{t-1,p-1}]\begin{bmatrix}\varphi^T(t-p)\\ H^T_{t-1,p-1}\end{bmatrix}=\varphi(t-p)\varphi^T(t-p)+P_\alpha^{-1}(t-1)$

$$\Rightarrow P_\alpha^{-1}(t-1)=P^{-1}(t-1)-\varphi(t-p)\varphi^T(t-p)$$

$$\therefore \alpha(t-1)=[I+P_\alpha(t-1)\varphi(t-p)\varphi^T(t-p)]\hat\theta(t-1)-P_\alpha(t-1)\varphi(t-p)y(t-p)$$

$$=\hat\theta(t-1)-P_\alpha(t-1)\varphi(t-p)[y(t-p)-\varphi^T(t-p)\hat\theta(t-1)]$$

有限数据窗的加权迭代最小二乘

（加权：$\hat{\theta}(t)=(H_t^T\Lambda H_t)^{-1}H_t^T\Lambda Y_t$）

$$p(t)=(H_t^T\Lambda H_t)^{-1}$$

$$p(t-1)=(H_{t-1}^T\Lambda H_{t-1})^{-1}$$

$$\begin{bmatrix}\lambda^{p-1} & & & \\ & \lambda^{p-2} & & \\ & & \ddots\lambda & \\ & & & 1\end{bmatrix}$$

$$\therefore\ \hat{\theta}(t)=p(t)H_t^T\Lambda Y_t=p(t)\sum_{i=t-p+1}^{t}\lambda^{t-i}\varphi(i)y(i)$$

$$=p(t)\left[\sum_{i=t-p}^{t-1}\lambda^{t-i}\varphi(i)y(i)+\varphi(t)y(t)-\lambda^p\varphi(t-p)y(t-p)\right]$$

$$=p(t)\left[p^{-1}(t-1)\,p(t-1)\sum_{i=t-p}^{t-1}\lambda^{t-i-1}\varphi(i)y(i)+\varphi(t)y(t)-\lambda^p\varphi(t-p)y(t-p)\right]$$

$$=p(t)p^{-1}(t-1)\,\lambda\hat{\theta}(t-1)+p(t)\varphi(t)y(t)-\lambda^p p(t)\varphi(t-p)y(t-p)$$

$$p^{-1}(t)=[\varphi(t-p+1)\ \cdots\ \varphi(t)]\begin{bmatrix}\lambda^{p-1} & & & \\ & \lambda^{p-2} & & \\ & & \ddots\lambda & \\ & & & 1\end{bmatrix}\begin{bmatrix}\varphi(t-p+1)\\ \vdots\\ \varphi^T(t-1)\\ \varphi^T(t)\end{bmatrix}=\sum_{i=t-p+1}^{t}\lambda^{t-i}\varphi(i)\varphi^T(i)$$

$$p^{-1}(t-1)=[\varphi(t-p)\ \cdots\ \varphi(t-1)]\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}\begin{bmatrix}\varphi^T(t-p)\\ \vdots\\ \varphi^T(t-1)\end{bmatrix}=\sum_{i=t-p}^{t-1}\lambda^{t-1-i}\varphi(i)\varphi^T(i)$$

有 $p^{-1}(t)=\lambda p^{-1}(t-1)+\varphi(t)\varphi^T(t)-\lambda^p\varphi(t-p)\varphi^T(t-p)$

$$\Rightarrow I=\lambda p(t)p^{-1}(t-1)+p(t)\varphi(t)\varphi^T(t)-\lambda^p p(t)\varphi(t-p)\varphi^T(t-p)$$

$$\hat{\theta}(t)=(I-p(t)\varphi(t)\varphi^T(t)+\lambda^p p(t)\varphi(t-p)\varphi^T(t-p))\hat{\theta}(t-1)+p(t)\varphi(t)y(t)$$
$$-\lambda^p p(t)\varphi(t-p)y(t-p)$$

$$\therefore\ \hat{\theta}(t)=\hat{\theta}(t-1)+p(t)\varphi(t)(y(t)-\varphi^T(t)\hat{\theta}(t-1))-\lambda^p p(t)\varphi(t-p)(y(t-p)-\varphi^T(t-p)\hat{\theta}(t-1))$$

$$\begin{cases}Q^{-1}(t)=\lambda P^{-1}(t-1)+\varphi(t)\varphi^T(t)\\ P^{-1}(t)=Q^{-1}(t)-\lambda^p\varphi(t-p)\varphi^T(t-p)\end{cases}$$

$$((A+BC)^{-1}=A^{-1}-A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1})$$

$$Q(t)=\frac{1}{\lambda}P(t-1)-\frac{1}{\lambda}P(t-1)\varphi(t)\frac{\varphi^T(t)\cdot\frac{1}{\lambda}P(t-1)}{I+\varphi^T(t)\frac{1}{\lambda}P(t-1)\varphi(t)}$$

$$=\frac{1}{\lambda}\left[P(t-1)-\frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{\lambda+\varphi^T(t)P(t-1)\varphi(t)}\right]$$

$$p(t)=Q(t)+\frac{Q(t)\lambda^p\varphi(t-p)\varphi^T(t-p)Q(t)}{1-\varphi^T(t-p)Q(t)\lambda^p\varphi(t-p)}=Q(t)+\frac{\lambda^pQ(t)\varphi(t-p)\varphi^T(t-p)Q(t)}{1-\lambda^p\varphi^T(t-p)Q(t)\varphi(t-p)}$$

求解迭代：$\alpha(t-1) = (H_{t-1,p-1}^T \Lambda H_{t-1,p-1})^{-1} H_{t-1,p-1}^T Y_{t-1,p-1}$

定义：

$$= [\varphi(t-p+1)\ \varphi(t-p+2) \cdots \varphi(t-1)] \begin{bmatrix} \lambda^{p-2} & & & \\ & \lambda^{p-3} & & \\ & & \ddots & \\ & & & \lambda \\ & & & & 1 \end{bmatrix} \begin{bmatrix} \varphi^T(t-p+1) \\ \varphi(t-p+1) \\ \vdots \\ \varphi(t-1) \end{bmatrix} \cdots$$

$P_\alpha(t-1)$

$$= \Big[\sum_{i=t-p+1}^{t-1} \lambda^{t-i-1} \varphi(i)\varphi^T(i)\Big]^{-1} \sum_{i=t-p+1}^{t-1} \lambda^{t-i-1} \varphi(i) y(i)$$

$P(t)$

$$\begin{cases} \hat\theta(t) = \Big[\sum_{i=t-p+1}^{t} \lambda^{t-i} \varphi(i)\varphi^T(i)\Big]^{-1} \sum_{i=t-p+1}^{t} \lambda^{t-i} \varphi(i) y(i) \\ \hat\theta(t-1) = \Big[\sum_{i=t-p}^{t-1} \lambda^{t-i-1} \varphi(i)\varphi^T(i)\Big]^{-1} \sum_{i=t-p}^{t-1} \lambda^{t-i-1} \varphi(i) y(i) \end{cases}$$

$P(t-1)$

$$\hat\theta(t) = P(t)\Big[\sum_{i=t-p+1}^{t-1} \lambda^{t-i} \varphi(i) y(i) + \varphi(t) y(t)\Big] = P(t)\Big[\lambda \sum_{i=t-p+1}^{t-1} \lambda^{t-i-1} \varphi(i) y(i) + \varphi(t) y(t)\Big]$$

$$= P(t)\Big[P_\alpha^{-1}(t-1)\ P_\alpha(t-1) \cdot \lambda \sum_{i=t-p+1}^{t-1} \lambda^{t-i-1} \varphi(i) y(t) + \varphi(t) y(t)\Big]$$

$$= \lambda P(t) P_\alpha^{-1}(t-1)\, \alpha(t-1) + P(t) \varphi(t) y(t)$$

$\Big(P^{-1}(t) = \lambda P_\alpha^{-1}(t-1) + \varphi(t)\varphi^T(t) \Rightarrow I = \lambda P(t) P_\alpha^{-1}(t-1) + P(t)\varphi(t)\varphi^T(t)\Big)$

$\therefore \hat\theta(t) = \alpha(t-1) + P(t)\varphi(t)\big(y(t) - \varphi^T(t)\alpha(t-1)\big)$

$$\hat\theta(t-1) = P(t-1) \sum_{i=t-p+1}^{t-1} \lambda^{t-i-1} \varphi(i) y(i) + \lambda^{p-1} \varphi(t-p) y(t-p)$$

$$= P(t-1)\Big[P_\alpha^{-1}(t-1)\ P_\alpha(t-1) \sum_{i=t-p+1}^{t-1} \lambda^{t-i+1} \varphi(i) y(i) + \lambda^{p-1} \varphi(t-p) y(t-p)\Big]$$

$$= P(t-1) P_\alpha^{-1}(t-1)\, \alpha(t-1) + \lambda^{p-1} P(t-1) \varphi(t-p) y(t-p)$$

$\Rightarrow P^{-1}(t-1)\, \theta(t-1) = P_\alpha^{-1}(t-1)\alpha(t-1) + \lambda^{p-1} \varphi(t-p) y(t-p)$

$\therefore \alpha(t-1) = P_\alpha(t-1) P^{-1}(t-1)\, \theta(t-1) - \lambda^{p-1} P_\alpha(t-1) \varphi(t-p) y(t-p)$

$\Big(P^{-1}(t-1) = P_\alpha^{-1}(t-1) + \lambda^{p-1} \varphi(t-p)\varphi^T(t-p) \Rightarrow P_\alpha(t-1) P^{-1}(t-1) = I + \lambda^{p-1} P_\alpha(t-1) \varphi(t-p)\varphi^T(t-p)\Big)$

$P_\alpha^{-1}(t-1) = P^{-1}(t-1) - \lambda^{p-1} \varphi(t-p)\varphi^T(t-p)$

$\therefore \alpha(t-1) = \theta(t-1) - \lambda^{p-1} P_\alpha(t-1) \varphi(t-p)\big(y(t-p) - \varphi^T(t-p)\theta(t-1)\big)$

考虑更复杂的噪声 (ARMAX模型)

$$\begin{cases} A(z^{-1})y(t)=B(z^{-1})u(t)+w(t) \\ w(t)=D(z^{-1})v(t) \end{cases}$$

$$A(z^{-1})=1+\sum_{i=1}^{n_a}a_iz^{-i}$$
$$B(z^{-1})=\sum_{i=1}^{n_b}b_iz^{-i}$$
$$D(z^{-1})=1+\sum_{i=1}^{n_d}d_iz^{-i}$$

差分 $y(t)=-a_1y(t-1)-\dots-a_{n_a}y(t-n_a)+b_1u(t-1)+\dots+b_{n_b}u(t-n_b)$
$+v(t)+d_1v(t-1)+\dots+d_{n_d}v(t-n_d)$

$$\hat{\varphi}(t)=\begin{bmatrix}-y(t-1)\\ \vdots\\ -y(t-n_a)\\ u(t-1)\\ \vdots\\ u(t-n_b)\\ \hat{v}(t-1)\\ \vdots\\ \hat{v}(t-n_d)\end{bmatrix},\ \theta=\begin{bmatrix}a_1\\ \vdots\\ a_{n_a}\\ b_1\\ \vdots\\ b_{n_b}\\ d_1\\ \vdots\\ d_{n_d}\end{bmatrix},\ y=\varphi^T(t)\theta+v(t)$$

$$\Longrightarrow\boxed{\begin{cases}\hat{\theta}(t)=\hat{\theta}(t-1)+L(t)[y(t)-\varphi^T(t)\hat{\theta}(t-1)]\\ L(t)=\dfrac{P(t-1)\varphi(t)}{1+\varphi^T(t)P(t-1)\varphi(t)}\\ P(t)=[I-L(t)\varphi^T(t)]P(t-1)\end{cases}}$$

问题：噪声不可测量. $\Rightarrow$ 估计噪声：在t时刻估计t-1时刻噪声

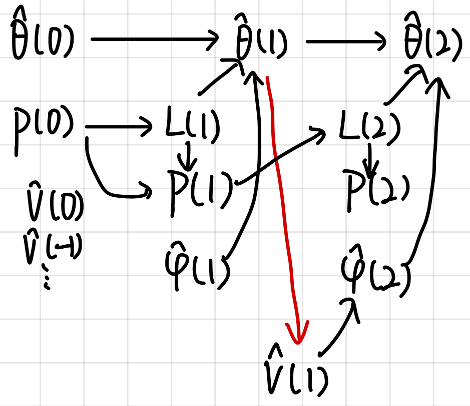
$$\begin{cases}v(t)=y(t)-\varphi^T(t)\theta\\ v(t-1)=y(t-1)-\varphi^T(t-1)\hat{\theta}(t-1)\end{cases}\Rightarrow\boxed{\begin{aligned}&\hat{v}(t)=y(t)-\varphi^T(t)\hat{\theta}(t)\\ &\hat{v}(t-1)=y(t-1)-\varphi^T(t-1)\hat{\theta}(t-1)\\ &\vdots\\ &\hat{v}(t-n_d)=y(t-n_d)-\varphi^T(t-n_d)\hat{\theta}(t-n_d)\end{aligned}}$$

$\Downarrow$ 改进

只用前一步的参数计算

$$\boxed{\begin{aligned}&\hat{v}(t)=y(t)-\varphi^T(t)\hat{\theta}(t)\\ &\hat{v}(t-1)=y(t-1)-\varphi^T(t)\hat{\theta}(t-1)\\ &\vdots\qquad\qquad\qquad\qquad \hat{\theta}(t-1)\\ &\hat{v}(t-n_d)=y(t-n_d)-\varphi^T(t)\hat{\theta}(t-1)\end{aligned}}$$

算法流程：



算法改进 $\Longrightarrow$

$$\theta=\begin{bmatrix}\theta_s\\ \theta_n\end{bmatrix},\ \theta_s=\begin{bmatrix}a_1\\ a_2\\ \vdots\\ a_{n_a}\\ b_1\\ \vdots\\ b_{n_b}\end{bmatrix},\ \theta_n=\begin{bmatrix}d_1\\ d_2\\ \vdots\\ d_{n_d}\end{bmatrix},\ 有\ y(t)=\varphi_s^T(t)\theta_s+\varphi_n^T(t)\theta_n+v(t)$$

focus on noise.

$$\varphi_s=\begin{bmatrix}-y(t-1)\\ \vdots\\ -y(t-n_a)\\ u(t-1)\\ u(t-n_b)\end{bmatrix},\ \varphi_n=\begin{bmatrix}v(t-1)\\ \vdots\\ v(t-n_d)\end{bmatrix}$$

令 $y_1(t)=y(t)-\varphi_s^T(t)\theta_s$：$y_1(t)=\varphi_n^T(t)\theta_n+v(t)$

$y_2(t)=y(t)-\varphi_n^T(t)\theta_n$：$y_2(t)=\varphi_s^T(t)\theta_s+v(t)$

$$\Rightarrow\theta_n:\begin{cases}\theta_n(t)=\theta_n(t-1)+L_n(t)[y_1(t)-\varphi_n^T(t)\hat{\theta}_n(t-1)]=\hat{\theta}_n(t-1)+L_n(t)[y(t)-\varphi_s^T(t)\hat{\theta}_s(t-1)-\hat{\varphi}_n^T(t)\hat{\theta}_n(t-1)]\\ L_n(t)=\dfrac{\hat{\varphi}_n(t)}{1+\hat{\varphi}_n^T(t)P_n(t-1)\hat{\varphi}_n(t)}P_n(t-1)\\ P_n(t)=[I-L_n(t)\hat{\varphi}_n(t)]P_n(t-1)\end{cases}$$

$y_1(t)$和θ_n有关，不可测

$$\theta_s = \begin{cases} \theta_s(t) = \theta_s(t-1) + L_s(t)[y_2(t) - \varphi_s^T(t)\hat{\theta}_s(t-1)] = \hat{\theta}_s(t-1) + L_s(t)[y(t) - \hat{\varphi}_n^T(t)\hat{\theta}_n(t) - \varphi_s^T(t)\hat{\theta}_s(t-1)] \\ L_s(t) = \dfrac{\varphi_s(t)}{1+\varphi_s^T(t)P_s(t-1)\varphi_s(t)} P_s(t-1) \\ P_s(t) = [I - L_s(t)\varphi_s(t)]P_s(t-1) \end{cases}$$

ARARX 模型： $A(z^{-1})y(t) = B(z^{-1})u(t) + \frac{1}{C(z^{-1})}v(t)$ $\left(C(z^{-1}) = 1 + \sum_{i=1}^{n_c} z^{-i}\right)$

$w(t) = \frac{1}{C(z^{-1})}v(t)$ 有 $\theta = \begin{bmatrix} a_1 \\ \vdots \\ a_{n_a} \\ b_1 \\ \vdots \\ b_{n_b} \\ c_1 \\ \vdots \\ c_{n_c} \end{bmatrix}$ $\varphi = \begin{bmatrix} -y(t-1) \\ \vdots \\ -y(t-n_a) \\ u(t-1) \\ \vdots \\ u(t-n_b) \\ -w(t-1) \\ \vdots \\ w(t-n_c) \end{bmatrix}$, $y(t) = \varphi^T(t)\theta + v(t)$

$w(t) = -c_1 w(t-1) - \dots - c w(t-n_c) + v(t)$

(不好直接转化成 ARX 模型，会增加参数且不易分解)

同样需要改进迭代（对 W 进行估计）

$\theta = \begin{bmatrix} \theta_s \\ \theta_n \end{bmatrix}$, $\varphi = \begin{bmatrix} \varphi_s \\ \varphi_n \end{bmatrix}$ $\therefore y(t) = \varphi_s^T(t)\theta_s + w(t)$ $\Rightarrow w(t) = y(t) - \varphi_s^T(t)\theta_s$

$\Downarrow$ 估计

$\hat{w}(t) = y(t) - \varphi_s^T(t)\hat{\theta}_s(t)$

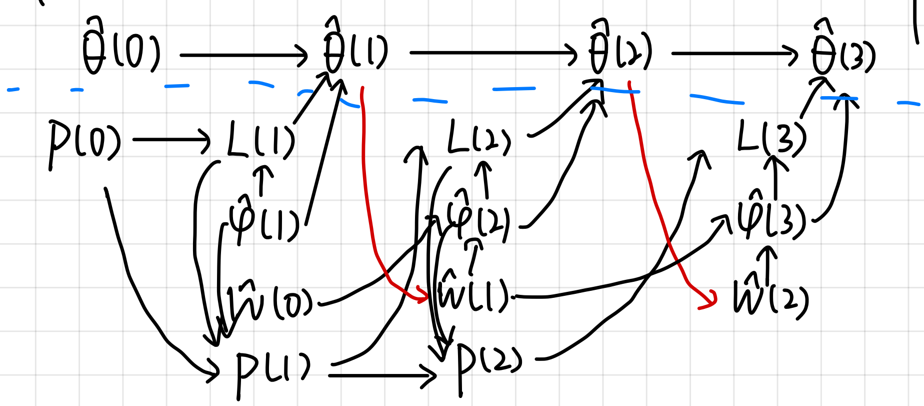
$$\Rightarrow \begin{cases} \hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)] \\ L(t) = \dfrac{P(t-1)\hat{\varphi}(t)}{1+\hat{\varphi}^T(t)P(t-1)\hat{\varphi}(t)} \\ P(t) = [I - L(t)\hat{\varphi}^T(t)]P(t-1) \end{cases}$$

$\hat{\varphi}(t) = \begin{bmatrix} \varphi_s(t) \\ \hat{\varphi}_n(t) \end{bmatrix}$

估计过程

$\hat{\varphi}_n(t) = \begin{bmatrix} \hat{w}(t-1) \\ \vdots \\ \hat{w}(t-n_c) \end{bmatrix}$ $\Rightarrow \begin{cases} \hat{w}(t-1) = y(t-1) - \varphi_s^T(t-1)\hat{\theta}_s(t-1) \\ \hat{w}(t-2) = y(t-2) - \varphi_s^T(t-2)\hat{\theta}_s(t-2) / \hat{\theta}_s(t-1) \\ \vdots \\ \hat{w}(t-n_c) = y(t-n_c) - \varphi_s^T(t-n_c)\hat{\theta}_s(t-n_c) / \hat{\theta}_s(t-1) \end{cases}$

算法流程：



$\Longrightarrow$ 处理 ARARX： $A(z^{-1})\underbrace{C(z^{-1})y(t)} = B(z^{-1})\underbrace{C(z^{-1})u(t)} + v(t)$

$\Longrightarrow A(z^{-1})y_f(t) = B(z^{-1})u_f(t) + v(t)$

$\boxed{\begin{aligned} &\theta_s = [a_1\ a_2 \dots a_{n_a}\ b_1 \dots b_{n_b}]^T \\ &\varphi_f(t) = [-y_f(t-1)\ -y_f(t-2) \dots -y_f(t-n_a)\ u_f(t-1) \dots u_f(t-n_b)]^T \end{aligned}}$ $\Rightarrow y_f(t) = \varphi_f^T(t)\theta_s + v(t)$

（转换成ARX模型形式）

$$\begin{cases} \hat{\theta}_s(t)=\hat{\theta}_s(t-1)+L_s(t)[\hat{y}_f(t)-\hat{\varphi}_f^T(t)\hat{\theta}_s(t-1)] \\ L_s(t)=\dfrac{P_s(t-1)\hat{\varphi}_f(t)}{1+\hat{\varphi}_f^T(t)P_s(t-1)\hat{\varphi}(t)} \\ P_s(t)=[I-L_s(t)\hat{\varphi}_f^T(t)]P_s(t-1) \end{cases}$$

(无法启动,需知 $\hat{y}_f(t)$, $\varphi_f^T(t)$

$$W(t)=\frac{1}{C(z^{-1})}V(t)\ ,\ C(z^{-1})W(t)=V(t)\Rightarrow \boxed{\begin{cases}\theta_n=[c_1\ c_2\ \dots\ c_{n_c}]^T \\ \varphi_n(t)=[-W(t-1)\ \cdots\ -W(t-n_c)]^T\end{cases}}$$

$$\Rightarrow W(t)=\varphi_n^T(t)\theta_n+V(t)$$

$y(t)-\varphi_s^T(t)\hat{\theta}_s(t-1)$

$\Downarrow$

$$\begin{cases} \hat{\theta}_n(t)=\hat{\theta}_n(t-1)+L_n(t)[\hat{W}(t)-\varphi_n^T(t)\hat{\theta}_n(t-1)] \\ L_n(t)=\dfrac{P_n(t-1)\varphi_n(t)}{1+\varphi_n^T(t)P_n(t-1)\varphi_n(t)} \\ P_n(t)=[I-L_n(t)\varphi_n^T(t)]P_n(t-1) \end{cases}$$

$\Rightarrow$ 估计 $\hat{\theta_n}\Rightarrow\hat{c}$

$$\hat{y}_f(t)=y(t)+\hat{c}_1(t-1)y(t-1)+\dots+\hat{c}_{n_c}(t-1)y(t-n_c)$$

$$\hat{u}_f(t)=u(t)+\hat{c}_1(t-1)u(t-1)+\dots+\hat{c}_{n_c}(t-1)u(t-n_c)$$

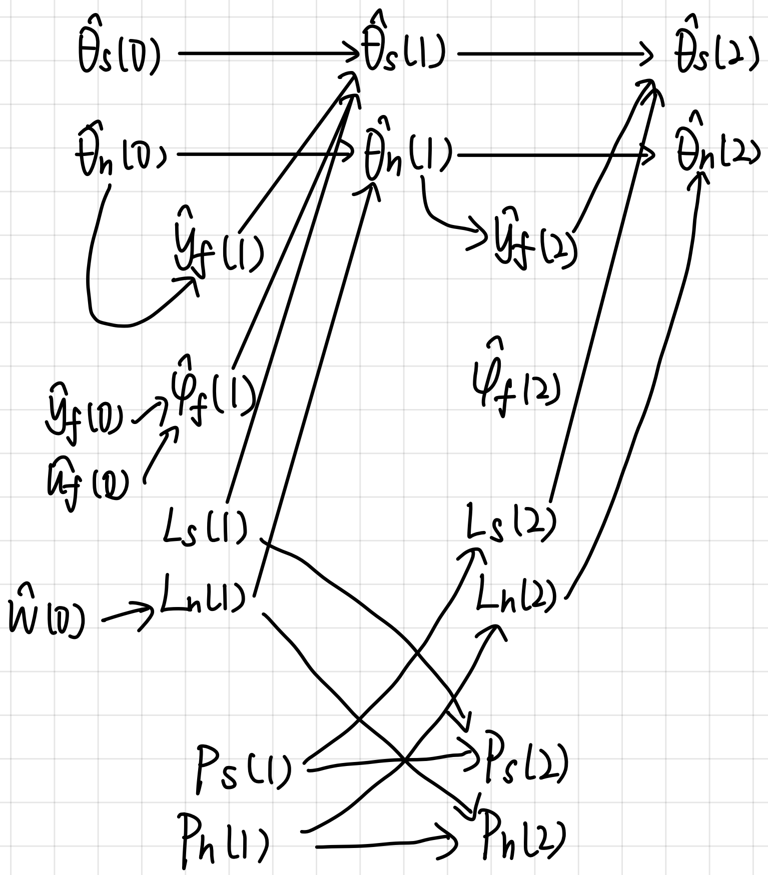
(无法启动需知 $\hat{W}(t)$)

$\downarrow$

$$A(z^{-1})y(t)=B(z^{-1})u(t)+W(t)$$

$\Rightarrow$ 辨识形式: $y(t)=\varphi_s^T(t)\theta_s+W(t)\Rightarrow\boxed{\hat{W}(t)=y(t)-\varphi_s^T(t)\hat{\theta}_s(t-1)}$

整理:



基于滤波器的递推
最小二乘辨识算法

优点:两步估计,
矩阵储存较小,降维

ARARMAX 模型辨识

方法一：$A(z^{-1})y(t) = B(z^{-1})u(t) + \frac{D(z^{-1})}{C(z^{-1})}v(t)$

$\Rightarrow A(z^{-1})y(t) = B(z^{-1})u(t) + w(t)$ $\qquad C(z^{-1})w(t) = D(z^{-1})v(t)$

$\theta_s = [a_1 \dots a_{n_a}\ b_1 \dots b_{n_b}]^T$

$\varphi_s(t) = [-y(t-1)\ -y(t-2) \dots -y(t-n_a)\ u(t-1)\ u(t-2) \dots u(t-n_b)]^T$

$\theta_n = [c_1 \dots c_{n_c}\ d_1 \dots d_{n_d}]$

$\varphi_n(t) = [-w(t-1) \dots -w(t-n_c)\ v(t-1) \dots v(t-n_d)]$

$y_t = \varphi_s^T(t)\theta_s + w(t)$ $\qquad w(t) = \varphi_n^T(t)\theta_n + v(t)$

$$y(t) = \varphi_s^T(t)\theta_s + \varphi_n^T(t)\theta_n + v(t)$$

令 $\theta = \begin{bmatrix}\theta_s \\ \theta_n\end{bmatrix}$，$\varphi(t) = \begin{bmatrix}\varphi_s(t) \\ \varphi_n(t)\end{bmatrix}$ $\therefore y(t) = \varphi(t)\theta + v(t)$

$$\Rightarrow \begin{cases} \hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \varphi^T(t)\hat{\theta}(t-1)] \\ L(t) = \frac{P(t-1)\varphi(t)}{1+\varphi^T(t)P(t-1)\varphi(t)} \\ P(t) = [I - L(t)\varphi(t)]P(t-1) \end{cases}$$

(不可能启动，φ中未知参数很多)
需对 $w(t)$，$v(t)$ 估计

估计：

$\hat{w}(t) = y(t) - \varphi_s^T(t)\hat{\theta}_s(t-1)$ $\qquad$ 可同上面基于滤波器的方法。

$\hat{v}(t) = y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1) = \hat{w}(t) - \hat{\varphi}_n^T(t)\hat{\theta}_n(t-1)$

$\qquad = y(t) - \varphi_s^T(t)\hat{\theta}_s(t-1) - \hat{\varphi}_n^T(t)\hat{\theta}_n(t-1)$

$\hat{\varphi}(t) = \begin{bmatrix}\varphi_s(t) \\ \hat{\varphi}_n(t)\end{bmatrix}$

$\Rightarrow \hat{\varphi}_n(t) = [-\hat{w}(t-1) \dots -\hat{w}(t-n_c)\ \hat{v}(t-1) \dots \hat{v}(t-n_d)]^T$

递阶的算法.

$y(t) - \varphi_s^T(t)\theta_s = \varphi_n^T(t)\theta_n + v(t)$

令 $y_1 = y(t) - \varphi_s^T(t)\theta_s$

$\therefore y_1(t) = \varphi_n^T(t)\theta_n + v(t)$ $\Rightarrow \hat{y}_1(t) = y(t) - \hat{\varphi}_s(t)\hat{\theta}_s(t-1)$

则有 RLS：

$$\begin{cases} \hat{\theta}_n(t) = \hat{\theta}_n(t-1) + L_n(t)[y(t) - \hat{\varphi}_s(t)\hat{\theta}_s(t-1) - \hat{\varphi}_n(t)\hat{\theta}_n(t-1)] \\ L_n(t) = \frac{P_n(t-1)\hat{\varphi}_n(t)}{1+\hat{\varphi}_n^T(t)P_n(t-1)\hat{\varphi}_n(t)} \\ P_n(t) = [I - L_n(t)\hat{\varphi}_n(t)]P_n(t-1) \end{cases}$$

$$y(t) - \varphi_n^T(t)\theta_n = \varphi_s^T(t)\theta_s + v(t)$$

令 $y_2 = y(t) - \varphi_n^T(t)\theta_n$

∴ $y_2 = \varphi_s^T(t)\theta_s + v(t) \Rightarrow \hat{y}_2 = y - \hat{\varphi}_n^T(t)\hat{\theta}_n(t-1)$

则有 RLS:

$$\begin{cases} \hat{\theta}_s(t) = \hat{\theta}_s(t-1) + L_s(t)[y(t) - \hat{\varphi}_n^T(t)\hat{\theta}_n(t-1) - \varphi_s(t)\hat{\theta}_s(t-1)] \\ L_s(t) = \dfrac{P_s(t-1)\varphi_s(t)}{1+\varphi_s^T(t)P_s(t-1)\varphi_s(t)} \\ P_s(t) = [I - L_n(t)\varphi_n(t)]P_n(t-1) \end{cases}$$

算法实现：先算 $\hat{\theta}_n(t)$，则计算 $\hat{\theta}_s(t)$ 时可用 $\underline{\hat{\theta}_n(t)}$ 代替 $\underline{\hat{\theta}_n(t-1)}$

filter-based recursive least square identification algorithms

基于滤波器的最小二乘算法.

[直接转化成方程误差模型]

方法二： $A(z^{-1})y(t) = B(z^{-1})u(t) + \dfrac{D(z^{-1})}{C(z^{-1})}v(t)$

$$A(z^{-1})\underbrace{\frac{C(z^{-1})}{D(z^{-1})}y(t)}_{y_f(t)} = B(z^{-1})\underbrace{\frac{C(z^{-1})}{D(z^{-1})}u(t)}_{u_f(t)} + v(t) \Rightarrow A(z^{-1})y_f(t) = B(z^{-1})u_f(t) + v(t)$$

∴ $\theta_s = [a_1 \dots a_{n_a}\ b_1 \dots b_{n_b}]$, $\varphi_f = [-y_f(t-1) \dots -y_f(t-n_a)\ u_f(t-1) \dots u_f(t-n_b)]^T$

∴ $y_f(t) = \varphi_f^T(t)\theta_s + v(t)$ [辨识形式]

$$\Rightarrow ① \begin{cases} \hat{\theta}_s(t) = \hat{\theta}_s(t-1) + L_s(t)[\hat{y}_f(t) - \hat{\varphi}_f^T(t)\hat{\theta}_s(t-1)] \\ L_s(t) = \dfrac{P_s(t-1)\hat{\varphi}_f(t)}{1+\hat{\varphi}_f^T(t)P_s(t-1)\hat{\varphi}_f(t)} \\ P_s(t) = [I - L_s(t)\hat{\varphi}_f^T(t)]P_s(t-1) \end{cases}$$

$$w(t) = \frac{D(z^{-1})}{C(z^{-1})}v(t) \Rightarrow C(z^{-1})w(t) = D(z^{-1})v(t)$$

∴ $\theta_n = [c_1 \dots c_{n_c}\ d_1 \dots d_{n_d}]$, $\hat{\varphi}_n(t) = [-\hat{w}(t-1) \dots -\hat{w}(t-n_c)\ \hat{v}(t-1) \dots \hat{v}(t-n_d)]^T$

∴ $w(t) = \varphi_n^T(t)\theta_n + v(t)$

$$\Rightarrow ② \begin{cases} \hat{\theta}_n(t) = \hat{\theta}_n(t-1) + L_n(t)[\hat{w}(t) - \hat{\varphi}_n^T(t)\hat{\theta}_n(t-1)] \\ L_n(t) = \dfrac{P_n(t)\hat{\varphi}_n(t)}{1+\hat{\varphi}_n^T(t)P_n(t)\hat{\varphi}_n(t)} \\ P_n(t) = [I - L_n(t)\hat{\varphi}_n^T(t)]P_n(t-1) \end{cases}$$

对 $w(t)$ 估计：$y(t) = \varphi_s^T(t)\theta_s + w(t)$　　$\therefore \hat{w}(t) = y(t) - \varphi_s^T(t)\hat{\theta}_s(t-1)$

可以代入式②中得到 $\hat{\theta}_n \Rightarrow \hat{c}, \hat{d}$　　$\therefore \hat{y}_f(t) = \frac{\hat{C}(z^{-1})}{\hat{D}(z^{-1})} y(t)$

$$\Rightarrow \hat{y}_f(t) = -\hat{d}_1(t-1)\hat{y}_f(t-1) - \dots - \hat{d}_{n_d}(t-1)\hat{y}_f(t-n_d) + y(t) + \hat{c}_1(t-1)y(t-1) + \dots + \hat{c}_{n_c}(t-1)y(t-n_c)$$

同理：$u_f(t) = \frac{C(z^{-1})}{D(z^{-1})} u(t)$

$$\Rightarrow \hat{u}_f(t) = -\hat{d}_1(t-1)\hat{u}_f(t-1) - \dots - \hat{d}_{n_d}(t-1)u_f(t-n_d) + u(t) + \hat{c}_1(t-1)u(t-1) + \dots + \hat{c}_{n_c}(t-1)u(t-n_c)$$

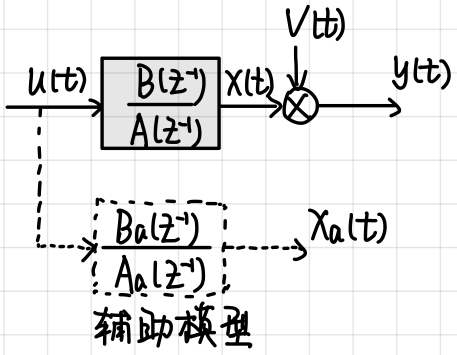
$\therefore$ 可得到 $\hat{\varphi}_f(t)$，改写式①

式②中仍需对 v 进行估计，从而得到 $\hat{\varphi}_n(t)$

有 $v(t) = w(t) - \hat{\varphi}_n^T(t)\theta_n \Longrightarrow \hat{v}(t) = \hat{w}(t) - \hat{\varphi}_n^T(t)\hat{\theta}_n(t-1)$

$$= y(t) - \varphi_s^T(t)\hat{\theta}_s(t-1) - \hat{\varphi}_n^T(t)\hat{\theta}_n(t-1)$$

最后一章　输出误差模型 $y(t)=\frac{B(z^{-1})}{A(z^{-1})}u(t)+w(t)$



考虑 $w(t)=v(t)$，即 $y(t)=\frac{B(z^{-1})}{A(z^{-1})}u(t)+v(t)$

有 $y(t)=x(t)+v(t)$

$$x(t)=\frac{B(z^{-1})}{A(z^{-1})}u(t)\Longrightarrow A(z^{-1})x(t)=B(z^{-1})u(t)$$

$\theta=[a_1\dots a_{n_a}\ b_1\dots b_{n_b}]^T$，$\varphi(t)=[-x(t-1)\dots -x(t-n_a)\ u(t-1)\dots u(t-n_b)]^T$

$\therefore x(t)=\varphi^T(t)\theta$

$\therefore y(t)=\varphi^T(t)\theta+v(t)$ [回归模型]

$\Longrightarrow$ 形式上 RLS：
$$\begin{cases}\hat{\theta}(t)=\hat{\theta}(t-1)+L(t)[y(t)-\varphi^T(t)\hat{\theta}(t-1)]\\ L(t)=\dfrac{P(t-1)\varphi(t)}{1+\varphi^T(t)P(t-1)\varphi(t)}\\ P(t)=[I-L(t)\varphi^T(t)]P(t-1)\end{cases}$$

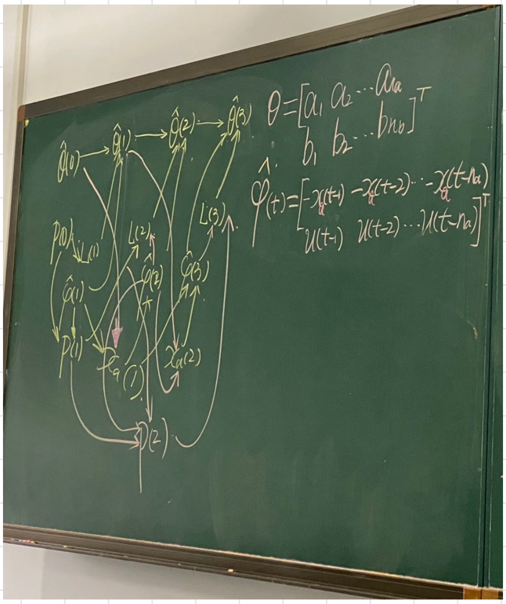
$$x_a(t)=\frac{\hat{b}_1(t-1)z^{-1}+\dots+\hat{b}_{n_b}(t-1)z^{-n_b}}{1+\hat{a}_1(t-1)z^{-1}+\dots+\hat{a}_{n_a}(t-1)z^{-n_b}}u(t)\Longrightarrow x_a(t)=-\hat{a}_1(t-1)x_a(t-1)\dots-\hat{a}_{n_a}(t-1)x_a(t-n_a)+\hat{b}_1(t-1)u(t-1)+\dots+\hat{b}_{n_b}(t-1)u(t-n_b)$$

$\hat{\varphi}(t)=[-x_a(t-1)\ \dots\ -x_a(t-n_a)\ \ u(t-1)\dots u(t-n_b)]$

$\therefore x_a(t)=\hat{\varphi}(t)\hat{\theta}(t-1)$

$\Longrightarrow$ RLS：
$$\begin{cases}\hat{\theta}(t)=\hat{\theta}(t-1)+L(t)[y(t)-\hat{\varphi}^T(t)\hat{\theta}(t-1)]\\ L(t)=\dfrac{P(t-1)\hat{\varphi}(t)}{1+\hat{\varphi}^T(t)P(t-1)\hat{\varphi}(t)}\\ P(t)=[I-L(t)\hat{\varphi}^T(t)]P(t-1)\end{cases}$$

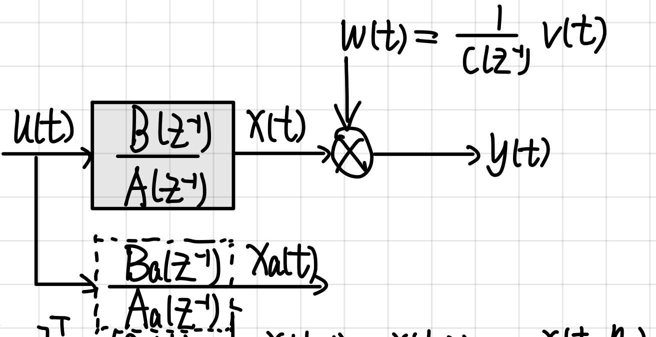
(两种模型处理本质都是变成辨识模型形式，输出误差model：把真实输出，



稍复杂模型：$y(t)=\frac{B(z^{-1})}{A(z^{-1})}u(t)+D(z^{-1})v(t)$

$$\frac{B(z^{-1})}{A(z^{-1})}u(t)=x(t) \quad , \quad \begin{aligned} y(t) &= x(t)+D(z^{-1})v(t) \\ &= x(t)+d_1 v(t-1)+d_2 v(t-2)+\cdots \\ &\quad + d_{n_d} v(t-n_d)+v(t) \\ &= \varphi_s^T(t)\theta_s+\varphi_n^T\theta_n+v(t) \end{aligned}$$

稍复杂模型：



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

$$\Longrightarrow y(t)=x(t)+w(t)$$

$$\begin{cases} x(t)=\frac{B(z^{-1})}{A(z^{-1})}u(t) \\ w(t)=\frac{1}{C(z^{-1})}v(t) \end{cases}$$

$\theta_s=[a_1 \dots a_{n_a}\ b_1 \dots b_{n_b}]^T$, $\varphi_s(t)=[-x(t-1)\ -x(t-2)\ \dots\ -x(t-n_a)\ u(t-1)\ \dots\dots\ u(t-n_b)]^T$

$\theta_n=[c_1\ c_2 \dots c_{n_c}]^T$, $\varphi_n(t)=[-w(t-1)\ \dots\ -w(t-n_c)]^T$

$\therefore x(t)=\varphi_s^T(t)\theta_s$, $w(t)=\varphi_n^T(t)\theta_n$

$\Longrightarrow y(t)=\varphi_s^T(t)\theta_s+\varphi_n^T(t)\theta_n+v(t)$ 令 $\theta=\begin{bmatrix}\theta_s\\ \theta_n\end{bmatrix}$, $\hat{\varphi}(t)=\begin{bmatrix}\hat{\varphi}_s(t)\\ \varphi_n(t)\end{bmatrix}$

RLS:
$$\begin{cases} \hat{\theta}(t)=\hat{\theta}(t-1)+L(t)[y(t)-\hat{\varphi}^T(t)\hat{\theta}(t-1)] \\ L(t)=\frac{P(t-1)\hat{\varphi}(t)}{1+\hat{\varphi}^T(t)P(t-1)\hat{\varphi}(t)} \\ P(t)=[I-L(t)\hat{\varphi}^T(t)]P(t-1) \end{cases}$$

无法启动，对 $\varphi(t)$ 估计 → 对 $x(t-1)\dots x(t-n_a)$ 估计 → 用辅助模型估计

$x_a(t)=\frac{B(z^{-1},t)}{\hat{A}(z^{-1},t)}u(t)$, $\begin{cases} A(z^{-1},t)=1+\hat{a}_1(t)z^{-1}+\dots+\hat{a}_{n_a}(t)z^{-n_a} \\ B(z^{-1},t)=\hat{b}_1(t)z^{-1}+\dots+\hat{b}_{n_b}(t)z^{-n_b} \end{cases}$

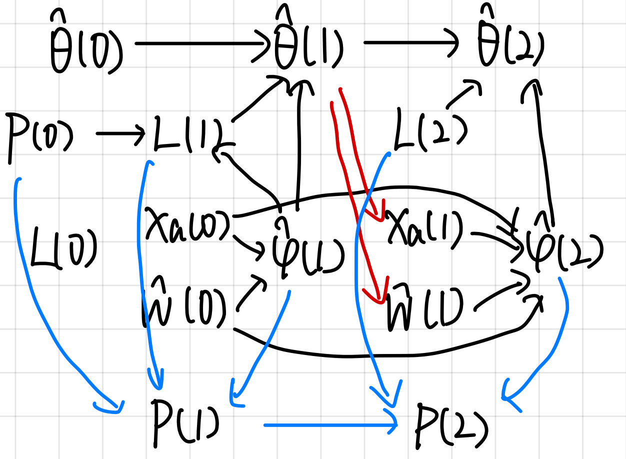
$\therefore$ 有 $x_a(t)=-\hat{a}_1(t)x_a(t-1)-\dots-\hat{a}_{n_a}(t)x_a(t-n_a)+\hat{b}_1(t)u(t-1)+\dots+\hat{b}_{n_b}(t)u(t-n_b)$

$\therefore \Longrightarrow \hat{\varphi}_s^T(t)=[-x_a(t-1)\ \dots\ -x_a(t-n_a)\ u(t-1)\dots u(t-n_b)]^T$ $\therefore x_a(t)=\hat{\varphi}_s^T(t)\theta_s(t)$

还需要对 $w(t)$ 进行估计：$y(t)=x(t)+w(t)$

$\therefore \hat{w}(t)=y(t)-x_a(t)=y(t)-\hat{\varphi}_s^T(t)\theta_s(t)$

分析算法：

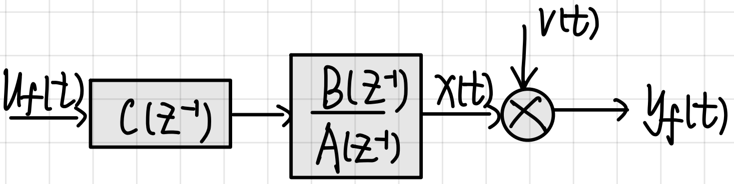


$\hat{\varphi}_n(t)=[-y(t-1)+\hat{\varphi}_s^T(t-1)\theta_s(t-1),$
$-y(t-2)+\hat{\varphi}_s^T(t-2)\theta_s(t-2)\dots]$

$\hat{\varphi}_s(t)=[-\varphi_s^T(t-1)\theta_s(t-1), -\varphi_s^T(t-2)\theta_s(t-2), \dots,$
$-\hat{\varphi}_s^T(t-n_a)\theta_s(t-n_a)]^T$

方法2：独立出白噪声

$$\underbrace{\boxed{C(z^{-1})y(t)}}_{y_f(t)}=\frac{B(z^{-1})}{A(z^{-1})}\underbrace{\boxed{C(z^{-1})u(t)}}_{u_f(t)}+v(t)\Rightarrow y_f(t)=\frac{B(z^{-1})}{A(z^{-1})}u_f(t)+v(t)$$



令 $x_f(t)=\frac{B(z^{-1})}{A(z^{-1})}u_f(t)$，$\theta_s=[a_1\dots a_{n_a}\ b_1\dots b_{n_b}]^T$

$\hat{\varphi}_f(t)=[-\hat{x}_f(t-1)\dots -\hat{x}_f(t-n_a)\ \hat{u}_f(t-1)\dots \hat{u}_f(t-n_b)]^T$

$\therefore y_f(t)=\varphi_f^T(t)\theta_s+v(t)$

辨识：
$$\begin{cases}\hat{\theta}_s(t)=\hat{\theta}_s(t-1)+L_s(t)(\hat{y}_f(t)-\hat{\varphi}_f^T(t)\theta_s(t-1))\\ L_s(t)=\dfrac{P_s(t-1)\hat{\varphi}_f(t)}{1+\hat{\varphi}_f^T(t)P_s(t-1)\hat{\varphi}_f(t)}\\ P_s(t)=[I-L_s(t)\hat{\varphi}_f^T(t)]P_s(t-1)\end{cases}$$

需要得到 u_f 和 y_f $\Rightarrow$ 对 $C(z^{-1})$ 估计 $\Rightarrow w(t)=\frac{1}{C(z^{-1})}v(t)$

$\therefore \varphi_n(t)=[-w(t-1)\dots -w(t-n_c)]^T$，$\theta_n=[c_1\dots c_{n_c}]$

$\therefore w(t)=\varphi_n^T(t)\theta_n+v(t)$

$$\Rightarrow\begin{cases}\hat{\theta}_n(t)=\hat{\theta}_n(t-1)+L_n(t)[\hat{w}(t)-\hat{\varphi}_n^T(t)\hat{\theta}_n(t-1)]\\ L_n(t)=\dfrac{P_n(t-1)\hat{\varphi}_n(t)}{1+\hat{\varphi}_n^T(t)P_n(t-1)\hat{\varphi}_n(t)}\\ P_n(t)=[I-L_n(t)\hat{\varphi}_n(t)]P_n(t-1)\end{cases}\qquad\Rightarrow\hat{\theta}_n\Rightarrow\hat{C}(z^{-1},t)$$

$\therefore y_f(t)=C(z^{-1})y(t)\Rightarrow \hat{y}_f(t)=\hat{C}(z^{-1},t)y(t)=y(t)+\varphi_y^T(t)\hat{\theta}_n(t-1)$

其中 $\varphi_y(t)=[y(t-1)\dots y(t-n_c)]^T$

$u_f(t)$：$\hat{u}_f(t)=u(t)+\varphi_u^T(t)\hat{\theta}_n(t-1)$

（有时间差）

$x_f(t)$：$\hat{x}_f(t)=\frac{\hat{B}(z^{-1},t)}{\hat{A}(z^{-1},t)}\hat{u}_f(t)\Rightarrow \hat{x}_f(t)=\hat{\varphi}_f^T(t)\hat{\theta}_s(t)$

$\hat{w}(t)$：$\hat{w}(t)=y(t)-\hat{x}(t)=y(t)-\varphi_a^T(t)\hat{\theta}_s(t)$

$\varphi_a(t)=[-x_a(t-1)\dots -x_a(t-n_a)\ u(t-1)\dots u(t-n_b)]^T$