

第四章：飞行器时域辨识方法

1. 基础知识
2. 最小二乘方法
3. 基于最小二乘方法的模型选择
4. 最大似然方法
5. 应用实例

随机变量

随机变量是指：

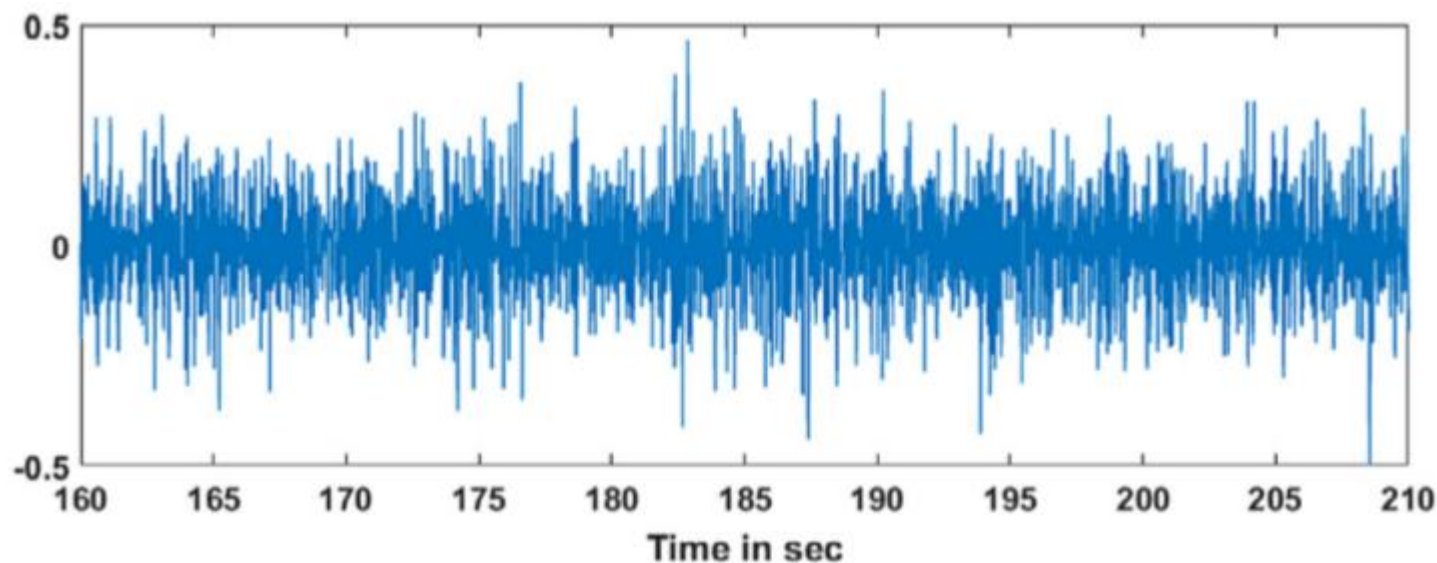
- 表示自然实值随机实验结果的变量
- 将随机实验的概率空间 (Ω, \mathcal{F}, P) 映射为实数的函数 X

在这两种情况下，实验的每一个可能结果都有一个实数

随机过程 - 实例

测量噪声是系统辨识中主要考虑的随机过程之一。

例如，加速度计静止时在某一轴上的测量数据是这样的：



概率分布

概率分布函数（PDF）（概率质量、概率密度）描述了随机变量取一定值的概率。

离散型概率分布

如果 $X: \Omega \rightarrow \mathbb{R}$ 为离散随机变量，其概率分布由其概率质量函数 $p_d: \mathbb{R} \rightarrow [0,1]$ 来描述。

随机变量取某值 x_i 的概率由下式给出：

$$P(X = x_i) = p_d(x_i) = P(\{\omega \in \Omega: X(\omega) = x_i\})$$

连续型概率分布

如果 $X: \Omega \rightarrow \mathbb{R}$ 为连续随机变量，其概率分布由其概率密度函数 $p(x): \mathbb{R} \rightarrow [0,1]$ 来描述。

随机变量落在给定区间 $[a, b]$ 内的概率由积分给出：

$$P(a \leq X \leq b) = \int_a^b p(x) dx$$

均值

考虑随机变量 X 有如下两种情况：

概率密度函数为 $p(x)$ 的连续情况；

概率质量函数为 $p_d(x)$ 的离散情况

X 的一阶矩被称为期望值、期望或均值，由下式给出：

$$E[X] = \sum_{i=1}^N x_i p_d(x) = \mu$$

或

$$E[X] = \int_{-\infty}^{\infty} xp(x)dx = \mu$$

方差

随机过程 X 的第二个中心矩是方差，它描述了数字与平均值（期望值）的距离。它是该变量与其平均值得平方偏差的期望值。

假设 X 的期望 $E[X] = \mu$ ，则变量 X 的方差可以写为：

$$\Rightarrow \text{Var}[X] = E[X^2] - (E[X])^2 = \sigma^2$$

其中， $\sigma = \sqrt{\text{Var}[X]}$ 为概率分布的标准差。用标准差代替方差描述分布的“散布”，其优点是标准差的单位与分布所描述的数据相同。

一、预备知识

正态分布

当一个随机变量 X 具有平均值 μ 和标准差 δ ，并且符合正态分布时， X 满足

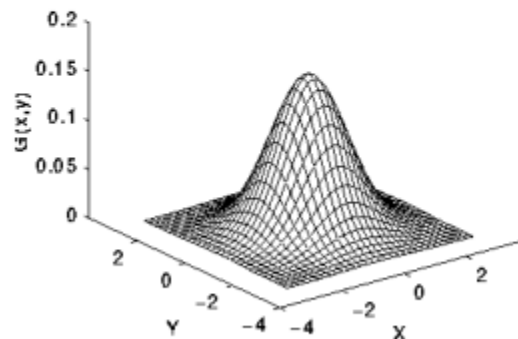
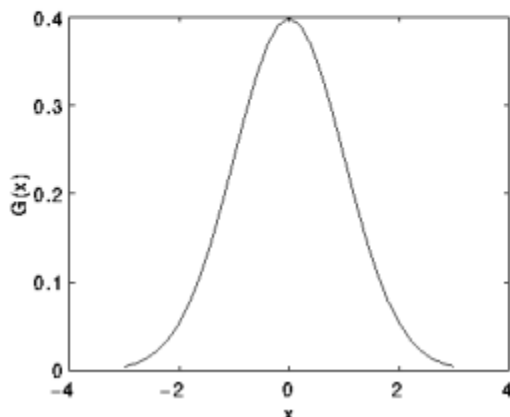
$$X \in N(\mu, \delta^2)$$

并且具有如下概率密度

$$p(x) = \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\delta^2}}$$

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$



一、预备知识

偏航力矩系数的模型:

$$C_n(i) = C_{n_0} + C_{n_\beta} \beta(i) + C_{n_p} \frac{b}{2V_0} p(i) + C_{n_r} \frac{b}{2V_0} r(i) + C_{n_{\delta_a}} \delta_a(i) + C_{n_{\delta_r}} \delta_r(i) + v_n(i)$$



根据最小二乘法:

输出变量 $y = \boxed{\theta_0} + \sum_{j=1}^n \boxed{\theta_j} \boxed{\xi_j}$ — $\beta \quad \frac{b}{2V_0} p \quad \frac{b}{2V_0} r \quad \delta_a \quad \delta_r$ **回归量**

$C_{n_0} \quad C_{n_\beta} \quad C_{n_p} \quad C_{n_r} \quad C_{n_{\delta_a}} \quad C_{n_{\delta_r}}$

回归方程 $\boxed{z(i)} = \theta_0 + \sum_{j=1}^n \theta_j \xi_j(i) + \boxed{v(i)} \quad i = 1, 2, \dots, N$

第*i*次量测输出

测量误差

N: 数据个数



一、预备知识

回归方程 $\boxed{z(i)} = \theta_0 + \sum_{j=1}^n \theta_j \xi_j(i) + \boxed{v(i)} \quad i = 1, 2, \dots, N$

假设：

- 模型参数 θ 是未知常数
- 输出量测误差服从正态分布 $N(0, \sigma)$

目标：基于最小二乘方法，估计 $\hat{\theta} = [\theta_0 \quad \theta_1 \quad \dots \quad \theta_n]$ 。

问题：

假设系统模型结构完全已知——灰盒模型，需要估计模型；
不同待估计参数可能线性相关——评价参数间的线性相关程度；
如何去评价估计结果的准确性。



二、最小二乘方法——2.1 基本方程

回归方程 $z = X\theta + v$

$z = [z(1) \ z(2) \ \cdots \ z(N)]^T = N \times 1$ 量测数列;

$\theta = [\theta_0 \ \theta_1 \ \cdots \ \theta_n]^T = n_p \times 1$ 未知参数数列;

$X = [1 \ \xi_1 \ \cdots \ \xi_n] = N \times n_p$ 数列, 由1和回归量构成的

$v = [v(1) \ v(2) \ \cdots \ v(N)]^T = N \times 1$ 测量噪声数列

$$J(\theta) = \frac{1}{2} (z - X\theta)^T (z - X\theta) = \frac{1}{2} \sum_{i=1}^N v_i^2 = \frac{1}{2} v^T v$$



$$\left\{ \begin{array}{l} \frac{\partial J}{\partial \theta} = -X^T z + X^T X \hat{\theta} = 0 \\ X^T X \hat{\theta} = X^T z \\ X^T (z - X \hat{\theta}) = 0 \end{array} \right.$$



$$\hat{\theta} = (X^T X)^{-1} X^T z \quad \text{估计值}$$

$$\theta = (X^T X)^{-1} X^T y \quad \text{真实值}$$



二、最小二乘方法——2.2 参数相关性分析

假设：每次测量都独立不相关，测量噪声间相互独立 $E(vv^T) = \sigma^2 I$

估计 $\hat{\theta}$ 的协方差矩阵 $\text{Cov}(\hat{\theta})$ ：

$$\text{Cov}(\hat{\theta}) = E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] = \sigma^2 (X^T X)^{-1}$$

定义矩阵 D

$$D = (X^T X)^{-1} = [d_{jk}] \quad j, k = 1, 2, \dots, n_p$$

第 j 个估计参数 $\hat{\theta}_j$ 的方差为协方差矩阵 $\text{Cov}(\hat{\theta})$ 的第 j 个主对角元。

$$\text{Var}(\hat{\theta}_j) = \sigma^2 d_{jj} = s^2(\hat{\theta}_j)$$

第 j 个估计参数 $\hat{\theta}_j$ 和第 k 个估计参数 $\hat{\theta}_k$ 之间的协方差为

$$\text{Cov}(\hat{\theta}_j, \hat{\theta}_k) = \sigma^2 d_{jk}$$



二、最小二乘方法——2.2 参数相关性分析

计算协方差矩阵 $\text{Cov}(\hat{\theta})$:

$$\text{Cov}(\hat{\theta}) = \text{E}[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] = \sigma^2 (X^T X)^{-1}$$

先验未知

对于重复性测量

$$\hat{\sigma}^2 = \frac{1}{(n_r - 1)} \sum_{i=1}^{n_r} [z_r(i) - \bar{z}_r]^2 = s^2$$

对于飞行试验 ,

1. s^2 可以通过试验数据平滑的方法得到 ;

2. s^2 可以通过残差分析得到

预测输出值

$$\hat{y} = X\hat{\theta} = X(X^T X)^{-1} X^T z$$

无偏估计

$$\hat{\sigma}^2 = \frac{v^T v}{(N - n_p)} = \frac{\sum_{i=1}^N [z(i) - \hat{y}(i)]^2}{(N - n_p)} = s^2$$

匹配误差

匹配误差：描述预测输出 \hat{y} 与量测输出 z 的接近程度。与模型结构、估计参数有关



二、最小二乘方法——2.2 参数相关性分析

第j个估计参数 $\hat{\theta}_j$ 和第k个估计参数 $\hat{\theta}_k$ 之间的**相关性系数** r_{jk} 为

$$r_{jk} = \frac{d_{jk}}{\sqrt{d_{jj}d_{kk}}} = \frac{\text{Cov}(\hat{\theta}_j, \hat{\theta}_k)}{\sqrt{\text{Var}(\hat{\theta}_j)\text{Var}(\hat{\theta}_k)}}$$

$$-1 \leq r_{jk} \leq 1$$

r_{jk} 是描述 $\hat{\theta}_j$ 和 $\hat{\theta}_k$ 之间的线性相关程度

$r_{jk} = 1$, $\hat{\theta}_j$ 和 $\hat{\theta}_k$ 线性相关 , 也表示相应的**回归量** ξ_j 和 ξ_k **线性相关**

参数自相关矩阵

$$\text{Corr}(\hat{\theta}) = [r_{jk}] \quad j, k = 1, 2, \dots, n_p$$



二、最小二乘方法——2.2 参数相关性分析

多项式各项间回归量 ξ_j 和 ξ_k 线性相关 \rightarrow 相关性系数 r_{jk} 接近1

求逆不准确



$X^T X$ 矩阵条件数大



$$\hat{\theta} = (X^T X)^{-1} X^T z$$

$$\text{Cov}(\hat{\theta}) = \text{E}[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] = \sigma^2 (X^T X)^{-1}$$

- 1). 协方差矩阵 $\text{Cov}(\hat{\theta})$ 主对角元素, $\text{Var}(\hat{\theta}_j)$ 越小越好
- 2). 自相关矩阵 $\text{Corr}(\hat{\theta})$ 非主对角元素 r_{jk} 越小越好

注意：飞行器参数辨识中，回归量常常线性相关



二、最小二乘方法——2.3 输出拟合性分析

除了估计残差方差 $\hat{\sigma}^2$ ，评价预测输出 \hat{y} 与量测输出 z 的接近程度，可以用**确定性系数** R^2

测量与其均值的偏差平方和 $SS_T = \sum_{i=1}^N [z(i) - \bar{z}]^2 = \mathbf{z}^T \mathbf{z} - N\bar{z}^2$ **仅与量测输出有关**

预测输出与测量均值的偏差平方和 $SS_R = \sum_{i=1}^N [\hat{y}(i) - \bar{z}]^2$ **模型/参数相关**

测量与预测输出的偏差平方和 $SS_E = \sum_{i=1}^N [z(i) - \hat{y}(i)]^2 = (\mathbf{z} - \mathbf{X}\hat{\boldsymbol{\theta}})^T (\mathbf{z} - \mathbf{X}\hat{\boldsymbol{\theta}})$
 $= \mathbf{z}^T \mathbf{z} - \hat{\boldsymbol{\theta}}^T \mathbf{X}^T \mathbf{z}$ **对于“好”的模型，仅与噪声相关**

$$SS_T = SS_R + SS_E \quad \longrightarrow \quad SS_R = SS_T - SS_E = \hat{\boldsymbol{\theta}}^T \mathbf{X}^T \mathbf{z} - N\bar{z}^2$$

确定性系数 R^2

$$R^2 = \frac{SS_R}{SS_T} = \frac{\hat{\boldsymbol{\theta}}^T \mathbf{X}^T \mathbf{z} - N\bar{z}^2}{\mathbf{z}^T \mathbf{z} - N\bar{z}^2} \quad 0 \leq R^2 \leq 1$$

$R^2 = 1$ 代表预测输出与测量输出的**完全拟合**。



二、最小二乘方法——2.4 残差分析

残差为 $v = z - \hat{y} = z - X\hat{\theta} = (I - X(X^T X)^{-1} X^T) z = (I - K)z$

$$K = X(X^T X)^{-1} X^T$$

期望 $E(v) = E(z) - KE(z) = X\theta - X(X^T X)^{-1} X^T X\theta = 0$

协方差 $\text{Cov}(v) = E(vv^T) = E[(I - K)zz^T(I - K)^T]$
 $= E[(v - Kv)(v - Kv)^T] = \sigma^2(I - K)$

对于第*i*次测量的残差 $v(i)$

$$\text{Var}[v(i)] = \sigma^2(I - \kappa_{ii})$$

$$\text{Cov}[v(i), v(j)] = -\sigma^2 \kappa_{ij}$$

当 $N \rightarrow \infty$, $\kappa_{ij} \rightarrow 0$, $\text{Cov}(v) \rightarrow \sigma^2 I$

因此，我们希望残差 v 接近白噪声特性  好的模型/参数

二、最小二乘方法——2.4 残差分析

残差白度分析：估计残差的自相关函数

$$\hat{R}_{vv}(k) = \frac{1}{N} \sum_{i=1}^{N-k} v(i)v(i+k) \quad k = 0, 1, \dots, N-1$$

当残差完全不相关， $\hat{R}_{vv}(k) = 0, k \neq 0$

实际上，即使对于合适的模型结构 $\hat{R}_{vv}(k) \neq 0$

好的白度： $\hat{R}_{vv}(k)$ 值在0附近，且要求应在 $-2s[\hat{R}_{vv}(k)] \leq \hat{R}_{vv}(k) \leq 2s[\hat{R}_{vv}(k)]$

残差的自相关函数估计的标准偏差为

$$s[\hat{R}_{vv}(k \neq 0)] \approx \frac{\hat{R}_{vv}(0)}{\sqrt{N}}$$

二、最小二乘方法——2.4 残差分析

残差白度分析：估计残差的自相关函数

$$\hat{R}_{vv}(k) = \frac{1}{N} \sum_{i=1}^{N-k} v(i)v(i+k) \quad k = 0, 1, \dots, N-1$$

当残差完全不相关， $\hat{R}_{vv}(k) = 0, k \neq 0$

实际上，即使对于合适的模型结构 $\hat{R}_{vv}(k) \neq 0$

好的白度： $\hat{R}_{vv}(k)$ 值在0附近，且要求应在 $-2s[\hat{R}_{vv}(k)] \leq \hat{R}_{vv}(k) \leq 2s[\hat{R}_{vv}(k)]$

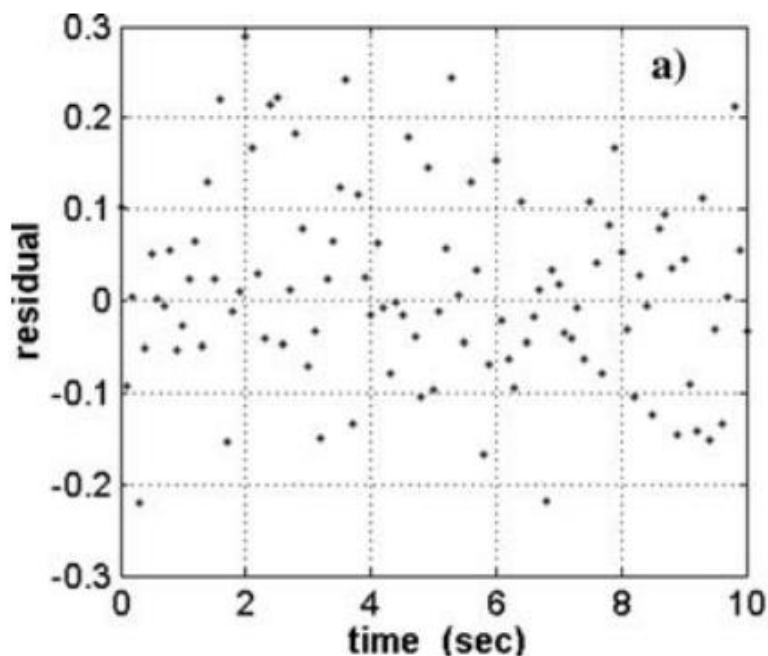
残差的自相关函数估计的标准偏差为

$$s[\hat{R}_{vv}(k \neq 0)] \approx \frac{\hat{R}_{vv}(0)}{\sqrt{N}}$$

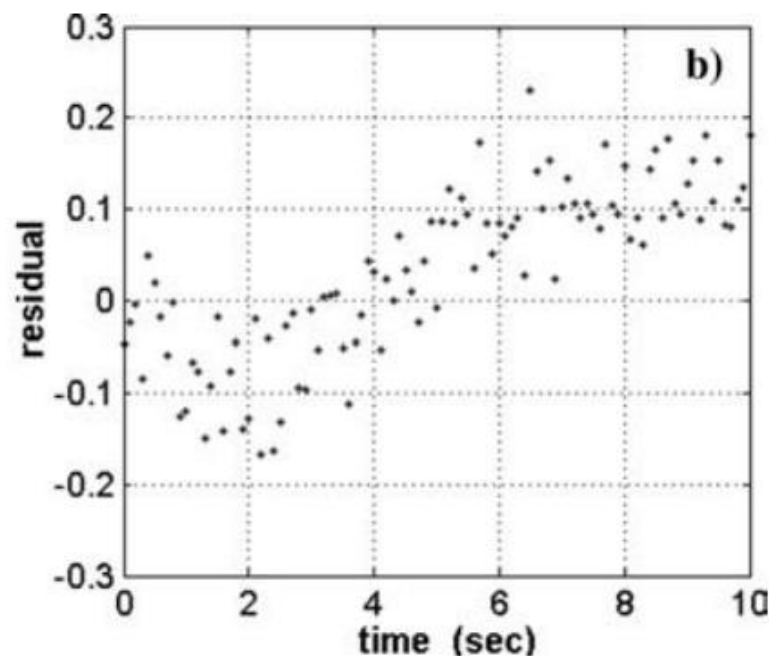
二、最小二乘方法——2.4 残差分析

残差检测：图像分析（可用于检测最小二乘法的准确性）

时域中的残差图像：被用来分析模型的准确度



随机的（接受）



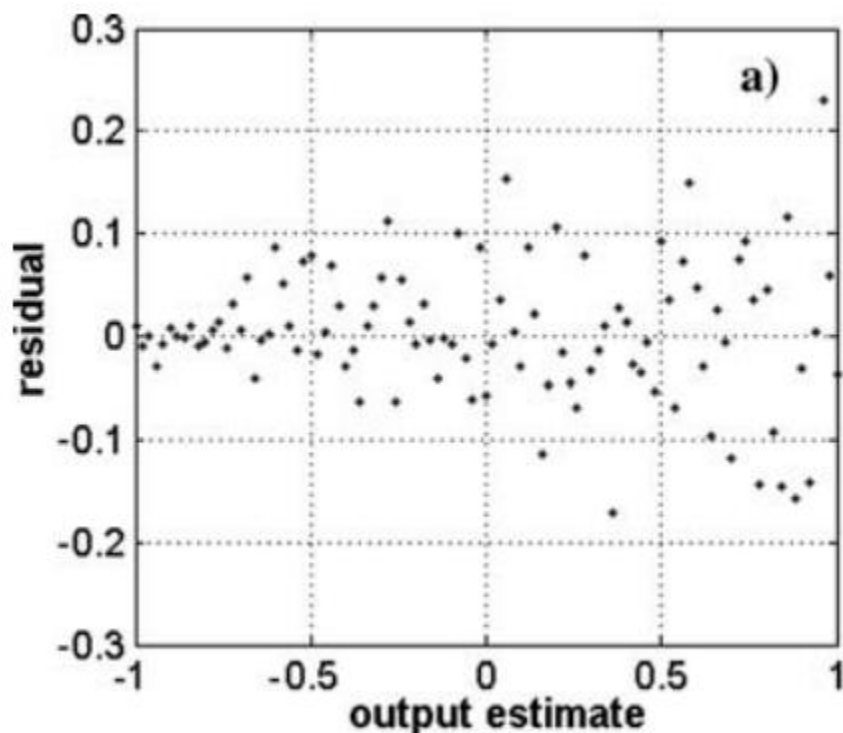
存在确定性成分（不接受）
模型缺失

二、最小二乘方法——2.4 残差分析

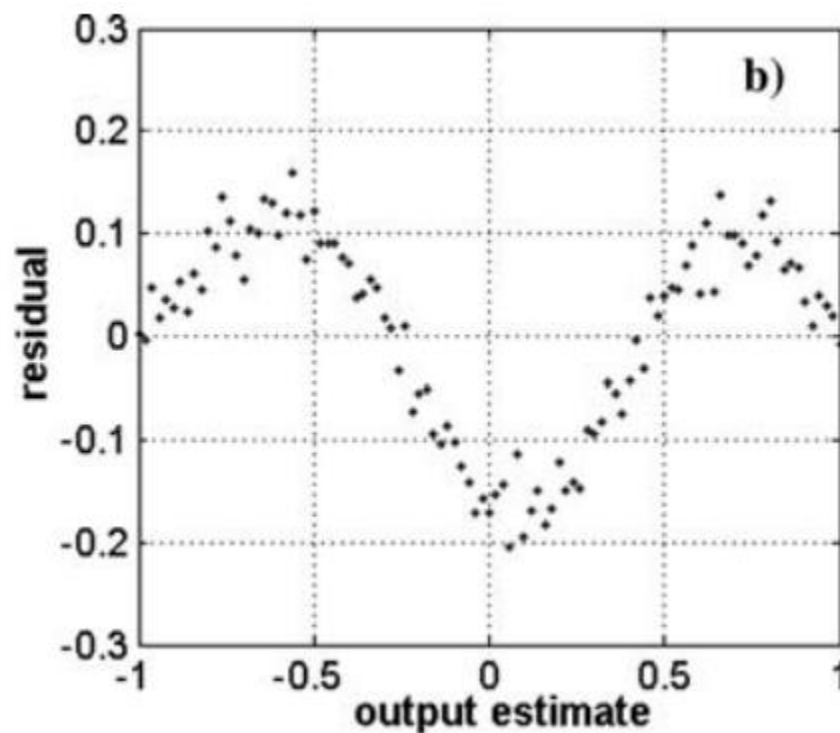


残差检测：图像分析（可用于检测最小二乘法的准确性）

残差随输出的变化：用于检测几种常见类型模型的不足之处。



残差方差不恒定，随输出变化



非线性特性没有被
模型表述



二、最小二乘方法——2.5 置信区间

置信区间：评价参数估计的质量

假设： $v \in \mathcal{N}(\mathbf{0}, \sigma^2 I)$

对于线性最小二乘估计器，

定理：服从高斯分布的变量的线性函数仍服从高斯分布
可得

$$\hat{\theta} \in \mathcal{N}\left(\mathbf{0}, \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}\right)$$

$$\hat{\mathbf{y}} \in \mathcal{N}\left(\mathbf{X}\theta, \sigma^2 \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T\right)$$



二、最小二乘方法——2.5 置信区间

理论基础：

1. 大数定理：

当样本量足够大时，这些样本 X_1 的均值 \bar{X}_1 无限接近总体的期望 $E(X)$ 。

2. 中心极限定理：

无论样本总体 X 满足什么样的部分，当样本数量足够大时，样本 X_1 的均值 \bar{X}_1 以正态分布的形式围绕总体均值 \bar{X} 波动，

$$\bar{X}_1 \in N(\bar{X}, s/\sqrt{N}) \quad s \text{ 为样本标准差}$$

二、最小二乘方法——2.5 置信区间

因此，知道总体标准差，可以利用正态分布评估样本均值的概率。

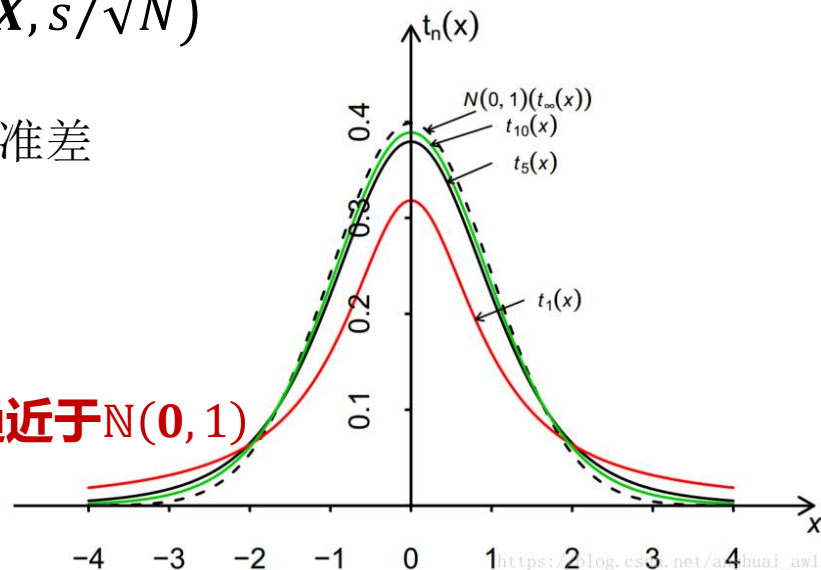
问题：样本值比较小，总体标准差未知

分布：
$$t = \frac{\bar{X}_1 - \bar{X}}{s/\sqrt{N}}$$
 s 为样本标准差

根据**中心极限定理**，只要样本量足够大，统计量的 X_1 的均值 \bar{X}_1 以正态分布的形式围绕总体均值 \bar{X} 波动， $\bar{X}_1 \in N(\bar{X}, s/\sqrt{N})$

t分布：
$$t = \frac{\bar{X}_1 - \bar{X}}{s/\sqrt{N}}$$
 s 为样本标准差

t分布特点：当样本量 N 趋于无穷大时，t分布趋近于 $N(0, 1)$



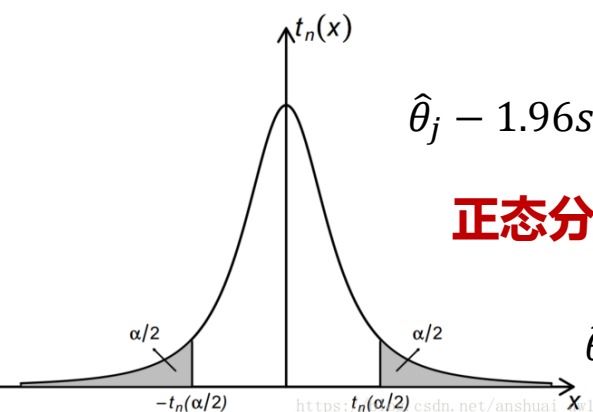
$\theta_j \sim t(N - n_p)$ 的***t***分布。

$$\mathbf{t} = \frac{\hat{\theta}_j - \theta_j}{\sqrt{\hat{\sigma}^2 d_{jj}}} = \frac{\hat{\theta}_j - \theta_j}{s(\hat{\theta}_j)} \quad j, k = 1, 2, \dots, n_p$$

$$\text{记 } \mathbf{P}\left(|\theta_j - \hat{\theta}_j| > cs(\hat{\theta}_j)\right) = \alpha$$

经验上：选取置信度95%，则 $\alpha = 0.05$,

$$\hat{\theta}_j - t\left(\frac{\alpha}{2}, N - n_p\right)s(\hat{\theta}_j) \leq \theta_j \leq \hat{\theta}_j + t(\alpha/2, N - n_p)s(\hat{\theta}_j)$$


$$\hat{\theta}_j - 1.96s(\hat{\theta}_j) \leq \theta_j \leq \hat{\theta}_j + 1.96s(\hat{\theta}_j)$$

正态分布

$$\hat{\theta}_j - 2s(\hat{\theta}_j) \leq \theta_j \leq \hat{\theta}_j + 2s(\hat{\theta}_j)$$

无人自主控制实验室
The institute of UAV autonomous control

二、最小二乘方法——2.5 置信区间



估计输出的 \hat{y} 的置信区间

$$\hat{y}(i) = \mathbf{X}^T(i)\hat{\boldsymbol{\theta}}$$

其方差为

$$\begin{aligned}\text{Var}[\hat{y}(i)] &= \mathbf{E}[(\hat{y}(i) - y(i))(\hat{y}(i) - y(i))^T] \\ &= \mathbf{E}\left[(\mathbf{X}^T(i)\hat{\boldsymbol{\theta}} - \mathbf{X}^T(i)\boldsymbol{\theta})(\mathbf{X}^T(i)\hat{\boldsymbol{\theta}} - \mathbf{X}^T(i)\boldsymbol{\theta})^T\right] \\ &= \mathbf{X}^T(i)\mathbf{E}\left[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\right]\mathbf{X}(i) \\ &= \sigma^2\mathbf{X}^T(i)(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}(i) = s^2[\hat{y}(i)]\end{aligned}$$

经验上：选取置信度95%，则 $\alpha = 0.05$,

$$\hat{y}(i) - t\left(\frac{\alpha}{2}, N - n_p\right)s(\hat{y}(i)) \leq y(i) \leq \hat{y}(i) + t(\alpha/2, N - n_p)s(\hat{y}(i))$$



二、最小二乘方法——2.5 置信区间

估计输出与量测误差（残差）的方差为

$$\begin{aligned}\text{Var}[z(i) - \hat{y}(i)] &= \mathbf{E}[(z(i) - \hat{y}(i))(z(i) - \hat{y}(i))] \\ &= \sigma^2 \left[1 + \mathbf{X}^T(i)(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}(i) \right] = s^2 [z(i) - \hat{y}(i)]\end{aligned}$$

可以看出

$$\text{Var}[v(i)] = \text{Var}[z(i) - \hat{y}(i)] = \boxed{\sigma^2} + \text{Var}[\hat{y}(i)]$$

测量误差方差 估计输出方差

经验上：选取置信度95%，则 $\alpha = 0.05$,

$$\hat{y}(i) - t\left(\frac{\alpha}{2}, N - n_p\right) s(z(i) - \hat{y}(i)) \leq z(i) \leq \hat{y}(i) + t(\alpha/2, N - n_p) s(z(i) - \hat{y}(i))$$

二、最小二乘方法——2.6 例1

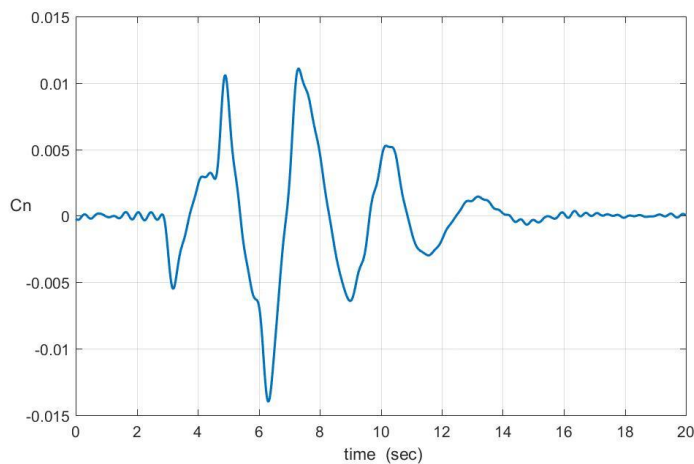
例子

偏航力矩系数的模型:

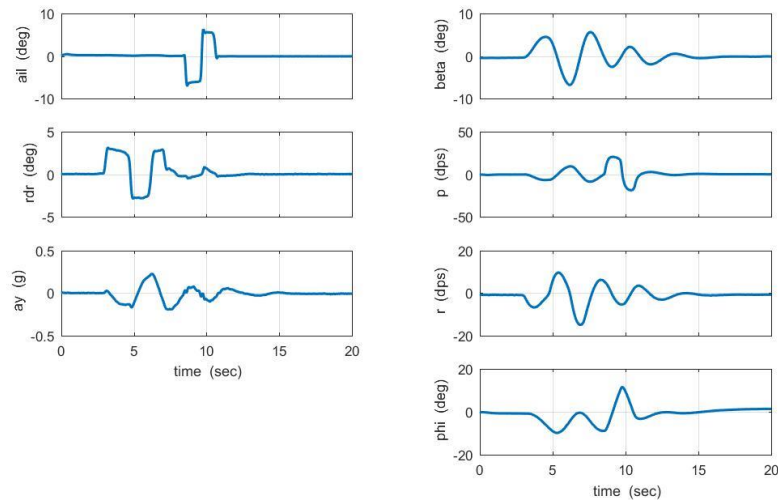
$$C_n(i) = C_{n0} + C_{n\beta}\beta(i) + C_{np}\frac{b}{2V_o}p(i) + C_{nr}\frac{b}{2V_o}r(i) + C_{n\delta_a}\delta_a(i) + C_{n\delta_r}\delta_r(i) + v_n(i)$$



测量得到的数据:



测量值z



各输入输出变量

二、最小二乘方法——2.6 例1

根据最小二乘方法：

$$\boldsymbol{\theta} = [C_{n_o} \ C_{n_\beta} \ C_{n_p} \ C_{n_r} \ C_{n_{\delta_a}} \ C_{n_{\delta_r}}]^T$$

$$\mathbf{z} = [C_n(1) \ C_n(2) \ \cdots \ C_n(N)]^T$$

$$\mathbf{X} = \begin{bmatrix} 1 & \beta(1) & \frac{b}{2V_o}p(1) & \frac{b}{2V_o}r(1) & \delta_a(1) & \delta_r(1) \\ 1 & \beta(1) & \frac{b}{2V_o}p(2) & \frac{b}{2V_o}r(2) & \delta_a(2) & \delta_r(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \beta(N) & \frac{b}{2V_o}p(N) & \frac{b}{2V_o}r(N) & \delta_a(N) & \delta_r(N) \end{bmatrix}$$

$$\hat{\boldsymbol{\theta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{z}$$

二、最小二乘方法——2.6 例1

$$\hat{\theta} = (X^T X)^{-1} X^T z$$

$$\text{Cov}(\hat{\theta}) = E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] = \sigma^2 (X^T X)^{-1}$$

$$\hat{\sigma}^2 = \frac{v^T v}{(N - n_p)} = \frac{\sum_{i=1}^N [z(i) - \hat{y}(i)]^2}{(N - n_p)}$$

$$t_0 = \frac{\hat{\theta}}{s(\hat{\theta})}$$

$C_{n_{\delta_a}}$: 副翼对偏航力矩的影响, 耦合项

Parameter	$\hat{\theta}$	$s(\hat{\theta})$	$ t_0 $	$100[s(\hat{\theta})/ \hat{\theta}]$
$C_{n_{\beta}}$	8.54×10^{-2}	3.58×10^{-4}	238.9	0.4
C_{n_p}	-5.15×10^{-2}	1.43×10^{-3}	35.9	2.8
C_{n_r}	-1.98×10^{-1}	1.30×10^{-3}	151.8	0.7
$C_{n_{\delta_a}}$	2.34×10^{-3}	5.00×10^{-4}	4.7	21.4
$C_{n_{\delta_r}}$	-1.31×10^{-1}	5.97×10^{-4}	218.5	0.5
C_{n_o}	-4.60×10^{-4}	7.42×10^{-6}	62.0	1.6
$s = \hat{\sigma}$	2.25×10^{-4}	—	—	—
$R^2, \%$	99.6	—	—	—

二、最小二乘方法——2.6 例1

第j个估计参数 $\hat{\theta}_j$ 和第k个估计参数 $\hat{\theta}_k$ 之间的**相关性系数** r_{jk} 为

$$r_{jk} = \frac{d_{jk}}{\sqrt{d_{jj}d_{kk}}} = \frac{\text{Cov}(\hat{\theta}_j, \hat{\theta}_k)}{\sqrt{\text{Var}(\hat{\theta}_j)\text{Var}(\hat{\theta}_k)}}$$

参数自相关矩阵 $\text{Corr}(\hat{\theta}) = [r_{jk}] \quad j, k = 1, 2, \dots, n_p$

1. 相关性系数都小于0.9； 2. $C_{n_{\delta a}}$ 和 C_{n_p} 都是耦合项，小量

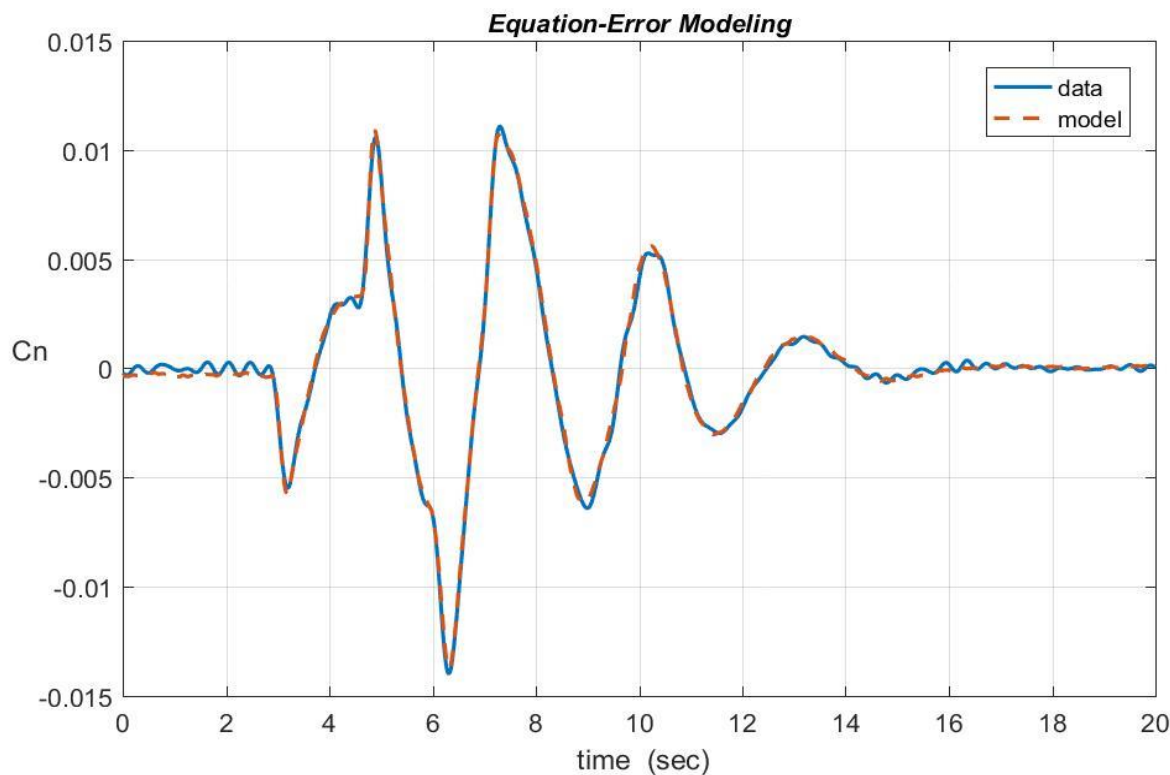
	C_{n_β}	C_{n_p}	C_{n_r}	$C_{n_{\delta a}}$	$C_{n_{\delta r}}$	C_{n_o}
C_{n_β}	1	0.82	0.15	0.67	-0.02	0.03
C_{n_p}	—	1	0.30	0.89	-0.18	-0.02
C_{n_r}	—	—	1	0.29	0.79	-0.21
$C_{n_{\delta a}}$	—	—	—	1	0.16	-0.01
$C_{n_{\delta r}}$	—	—	—	—	1	-0.28
C_{n_o}	—	—	—	—	—	1

二、最小二乘方法——2.6 例1

预测输出值 $\hat{y} = X\hat{\theta} = X(X^T X)^{-1} X^T z$

$$\text{Var}[\hat{y}(i)] = \sigma^2 X^T(i)(X^T X)^{-1} X(i) = s^2[\hat{y}(i)]$$

95%置信区间 $t\left(\frac{0.05}{2}, N - n_p\right) s(\hat{y}(i)) = 3 \times 10^{-5}$



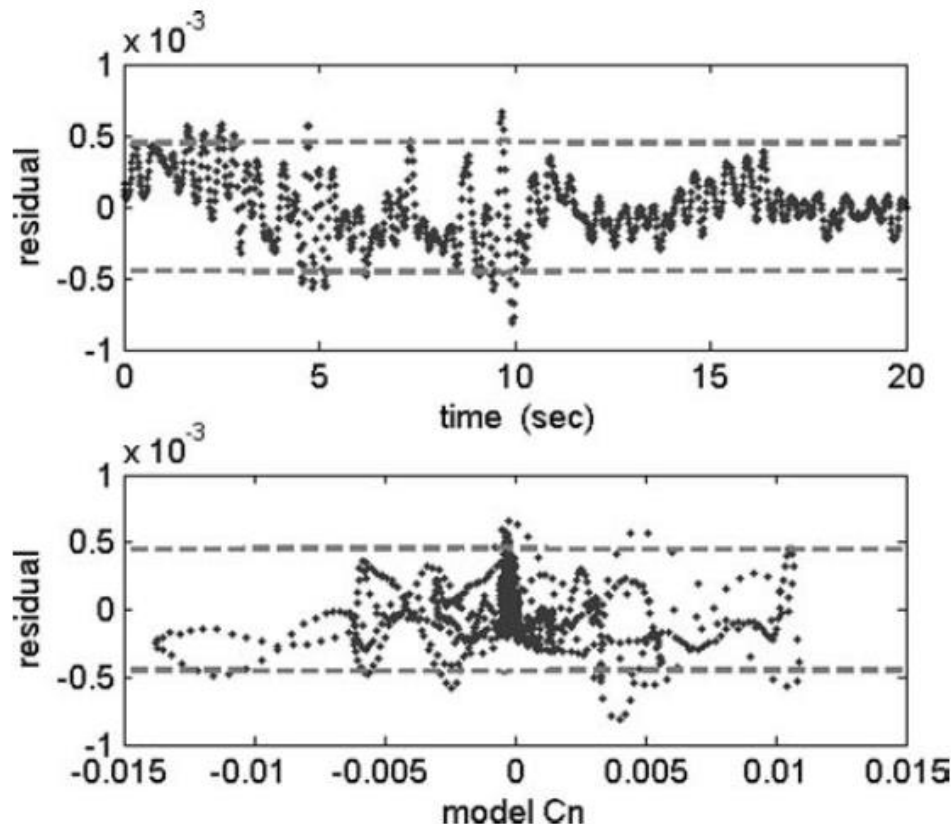
二、最小二乘方法——2.6 例1

估计输出与量测误差（残差）的方差为

$$\text{Var}[v(i)] = \sigma^2 \left[1 + \mathbf{X}^T(i)(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}(i) \right]$$

经验上：选取置信度95%，则 $\alpha = 0.05$,

$$t\left(\frac{0.05}{2}, N - n_p\right) s(z(i) - \hat{y}(i)) = 4.5 \times 10^{-4}$$



二、最小二乘方法——2.6 例1

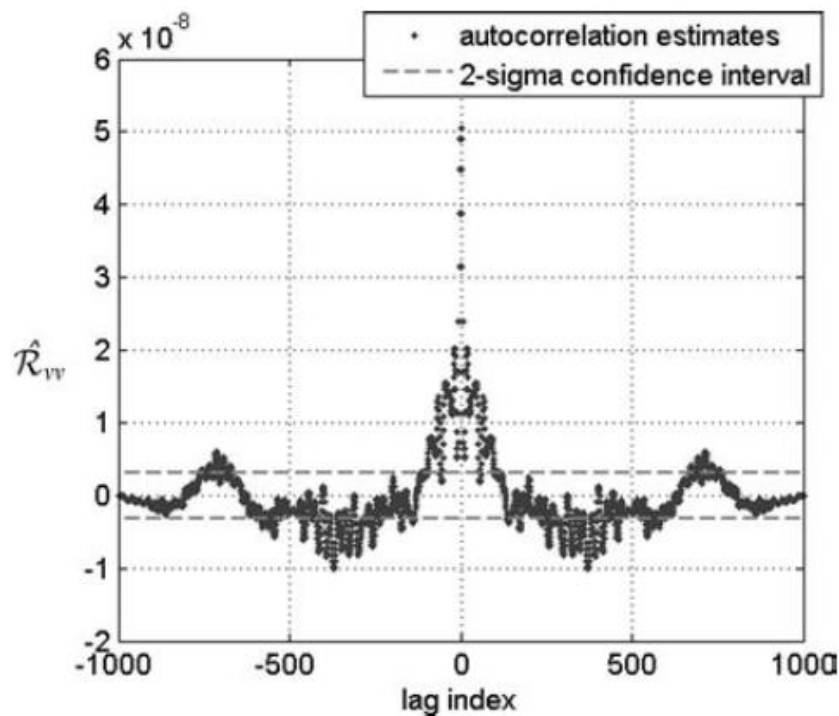
残差自相关函数为

$$\hat{R}_{vv}(k) = \frac{1}{N} \sum_{i=1}^{N-k} v(i)v(i+k) \quad k = 0, 1, \dots, N-1$$

估计标准差 $s[\hat{R}_{vv}(k \neq 0)] \approx \frac{\hat{R}_{vv}(0)}{\sqrt{N}}$

95%置信度

$$-2s[\hat{R}_{vv}(k)] \leq \hat{R}_{vv}(k) \leq 2s[\hat{R}_{vv}(k)]$$



二、最小二乘方法——2.6 例1

残差自相关函数为

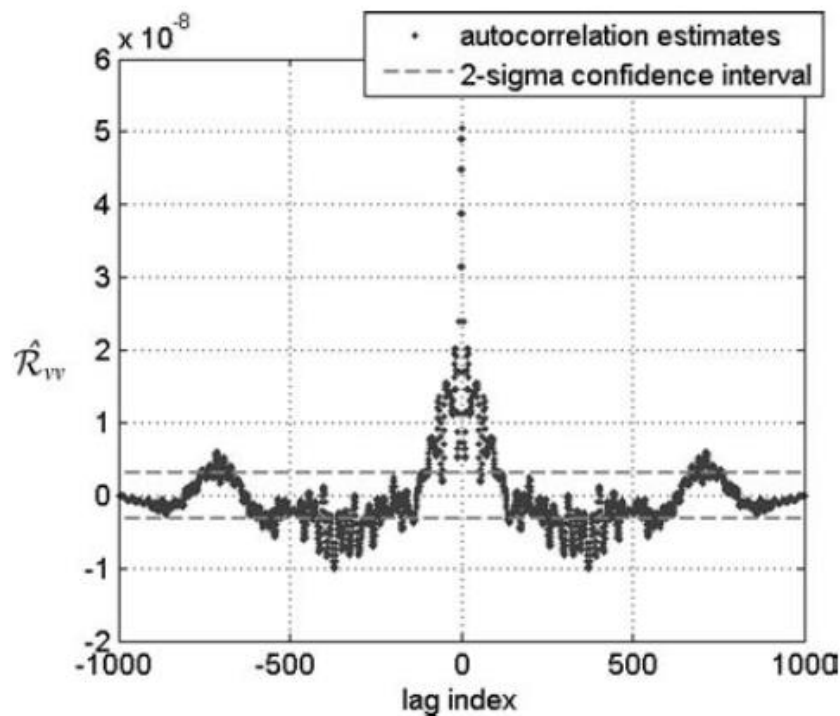
$$\hat{R}_{vv}(k) = \frac{1}{N} \sum_{i=1}^{N-k} v(i)v(i+k) \quad k = 0, 1, \dots, N-1$$

估计标准差 $s[\hat{R}_{vv}(k \neq 0)] \approx \frac{\hat{R}_{vv}(0)}{\sqrt{N}}$

95%置信度

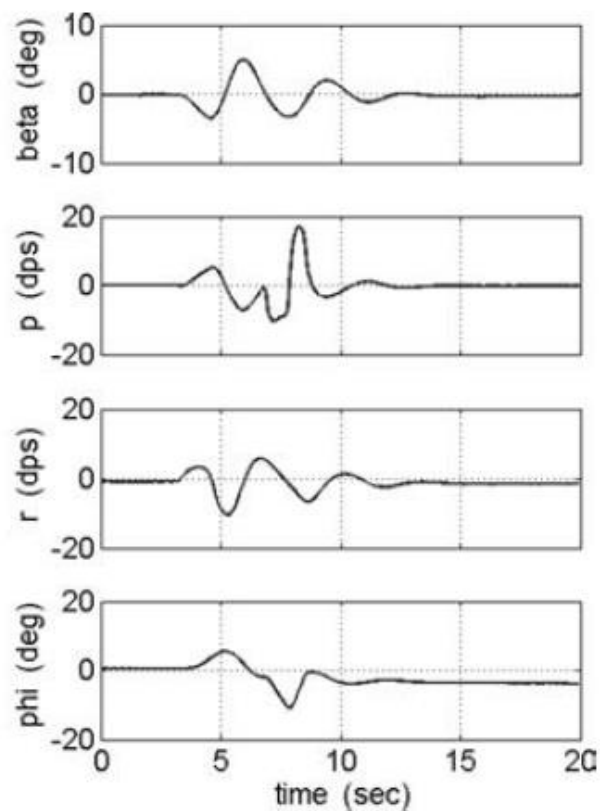
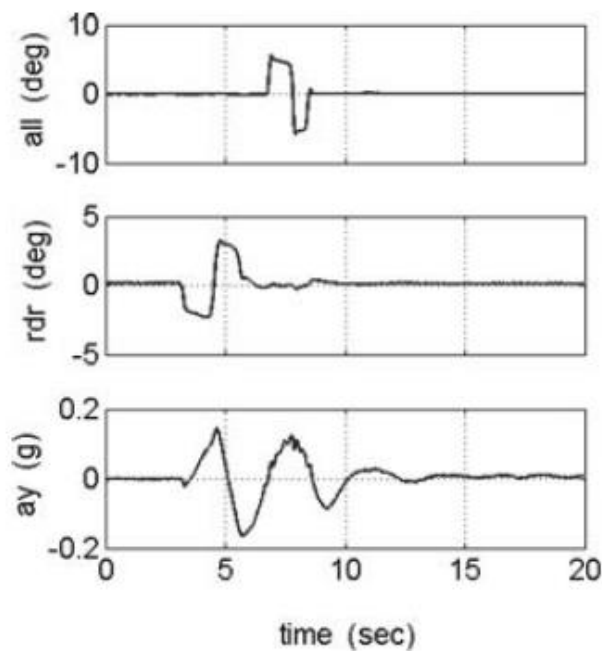
$$-2s[\hat{R}_{vv}(k)] \leq \hat{R}_{vv}(k) \leq 2s[\hat{R}_{vv}(k)]$$

若为白噪声, $k = 0$, 处为峰值 $\hat{R}_{vv}(0)$
 $k \neq 0$ 时, $\hat{R}_{vv}(k) \rightarrow 0$



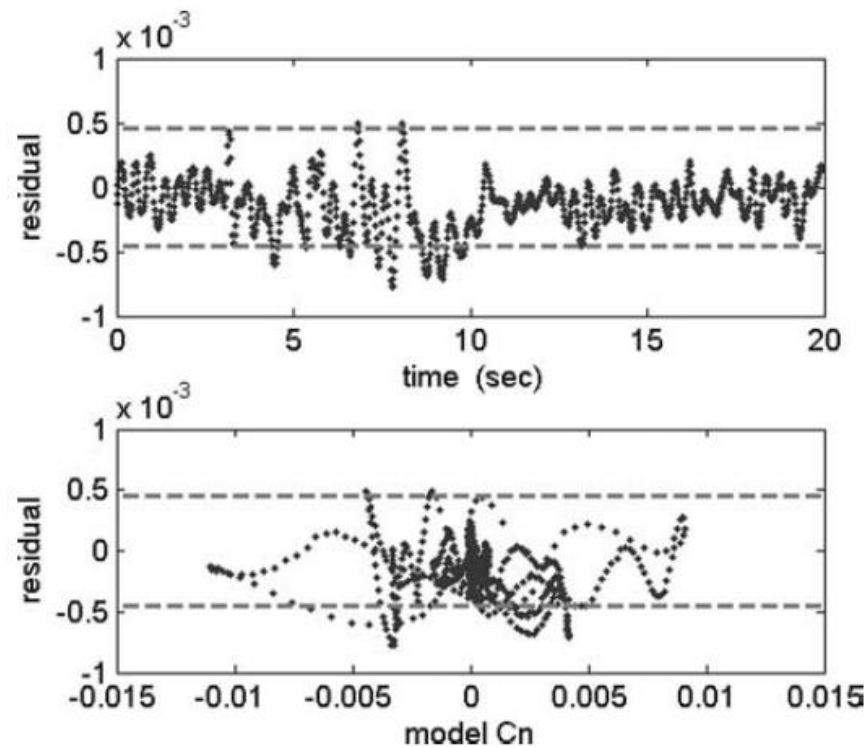
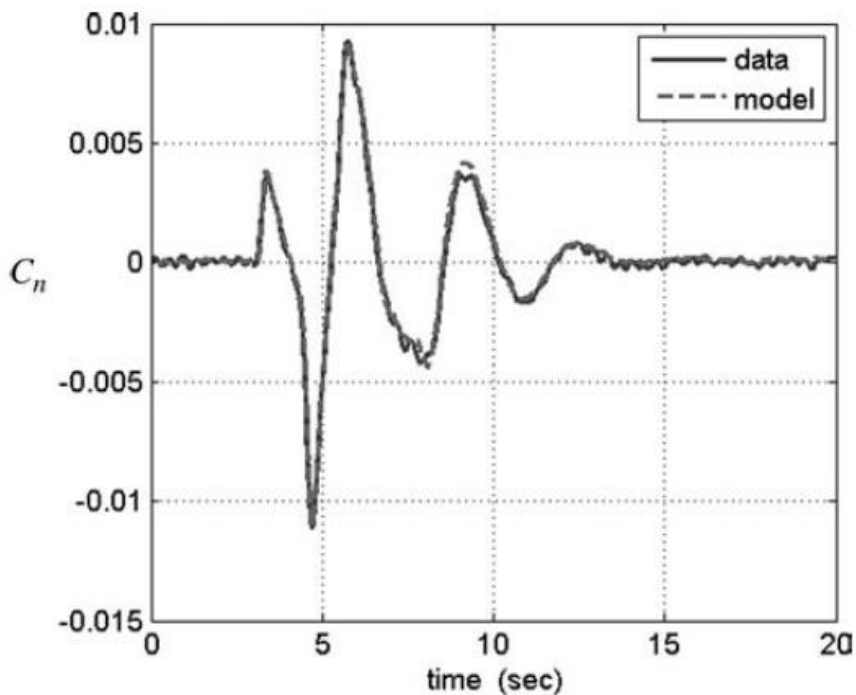
二、最小二乘方法——2.6 例1

模型检验



二、最小二乘方法——2.6 例1

模型检验



三、模型选择



模型结构的选择

要点:


- 1.控制上希望模型更为简单, 这样设计更容易实现, 也具有更广泛的应用
- 2.然而对于某个系统或者过程, 太过简单的模型很难充分描述, 模型越复杂, 对系统的描述更完备



吝啬原则: 给定两个模型, 如果它们拟合相同的数据且残差方差几乎相等, 选择参数最少的模型

三、模型选择

模型结构的选择

$$C_n(i) = C_{n0} + C_{n\beta}\overset{\text{侧滑角}}{\beta(i)} + C_{np}\frac{b}{2V_o}\overset{\text{滚转率}}{p(i)} + C_{nr}\frac{b}{2V_o}\overset{\text{偏航率}}{r(i)} \\ + C_{n\delta_a}\overset{\text{副翼输入}}{\delta_a(i)} + C_{n\delta_r}\overset{\text{舵偏}}{\delta_r(i)} + v_n(i)$$


与偏航力矩系数相关的参数非常多，为何选择上述参数？

选择参数存在的矛盾点：

一方面，可能需要具有许多回归变量的模型，以便该模型可以描述因变量的几乎所有变化。

在另一方面，我们希望有一个模型，回归模型越少越好，因为预测的方差随着回归器数量的增加而增加增加



精简模型的性质

模型为

$$z = X\theta + v$$

也可以表示为

$$z = X_1\theta_1 + X_2\theta_2 + v$$

当去除 θ_2 中对应的一个或者多个参数，得到一个精简模型：

$$z = X_1\theta_1 + v$$

结果：删除小值且具有较大标准差的参数能提高精度，但是删除参数会导致偏差。所以 $\tilde{\theta}_1$ 是一个有偏估计，原本的 $\hat{\theta}_1$ 是无偏估计

同时有如下性质

$$Var(\tilde{\theta}_1) \leq Var(\hat{\theta}_1)$$

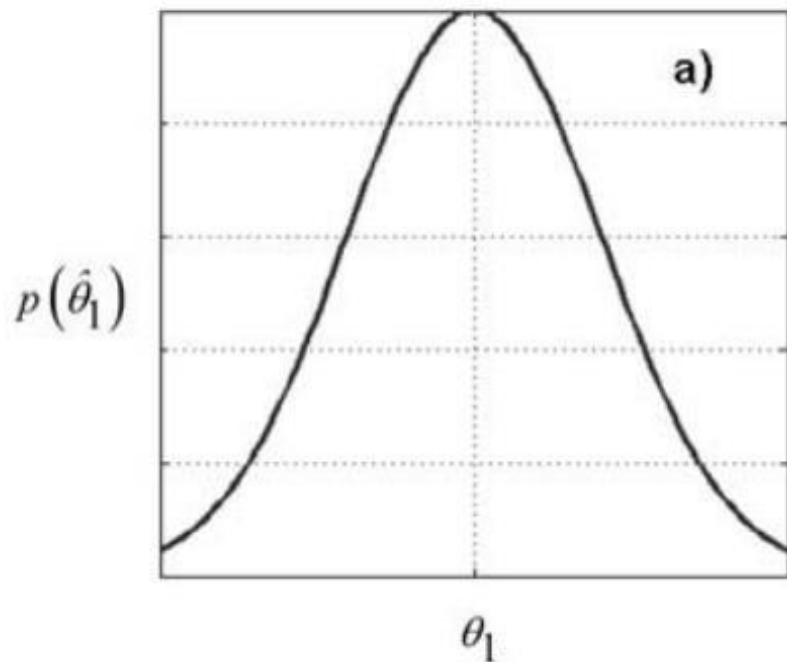
$$MSE(\tilde{\theta}_1) \leq Var(\hat{\theta}_1)$$

其中，均方误差为 $MSE(\tilde{\theta}_1) = Var(\tilde{\theta}_1) + [E(\tilde{\theta}_1) - \theta_1]^2$

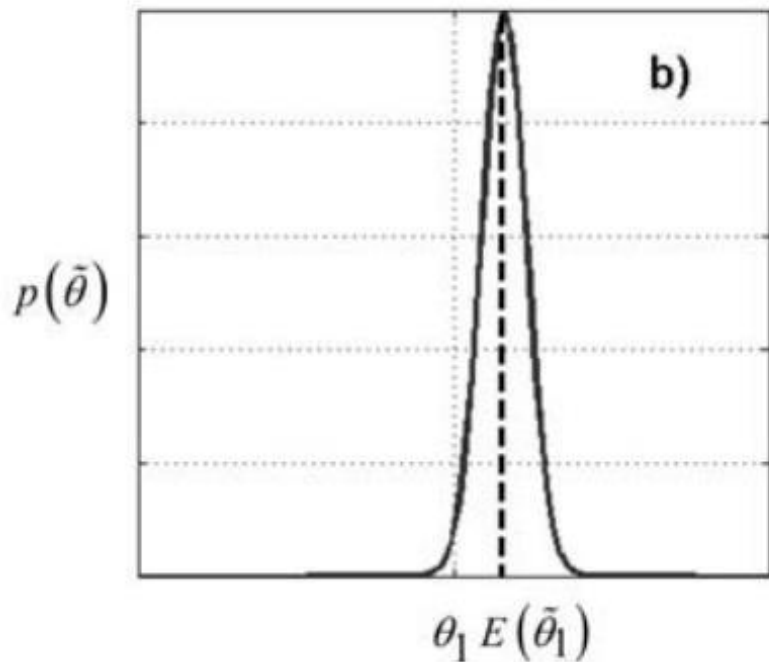
三、模型选择



$$MSE(\hat{\theta}_1) = Var(\hat{\theta}_1)$$



$$MSE(\tilde{\theta}_1) = Var(\tilde{\theta}_1) + [E(\tilde{\theta}_1) - \theta_1]^2$$



概率密度示意图

(要点: 图形宽度和 θ 和 $E(\theta)$ 的差别)

三、模型选择

$$\hat{\sigma}^2 = \frac{\mathbf{v}^T \mathbf{v}}{(N - n_p)} = \frac{1}{(N - n_p)} \sum_{i=1}^N [z(i) - \hat{y}(i)]^2$$

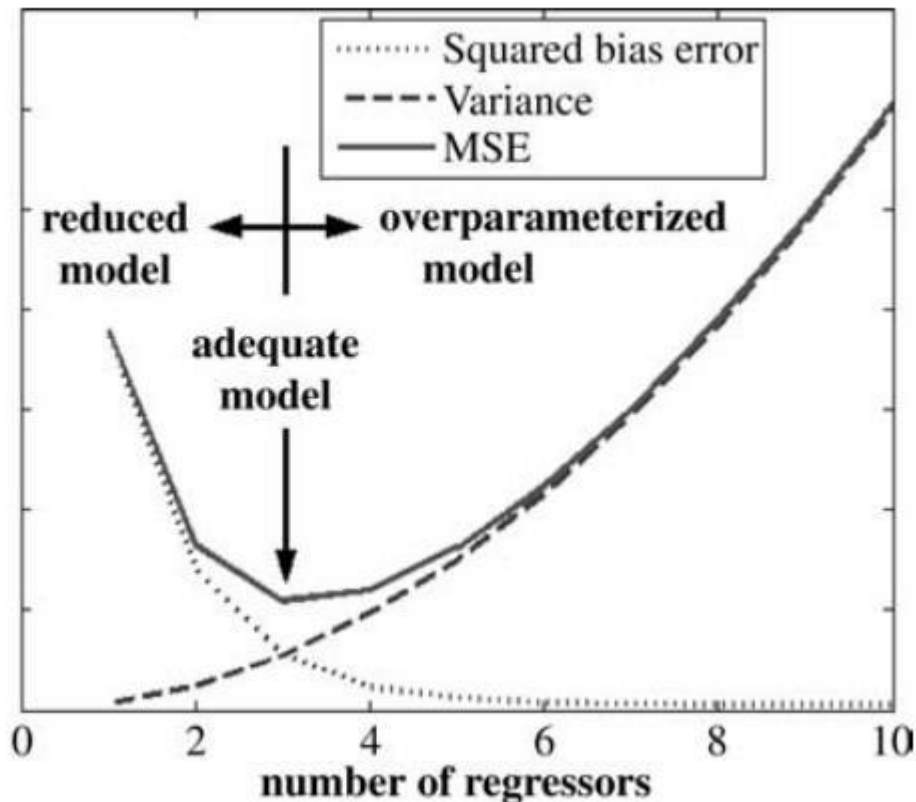
方差表示估计整体偏离均值的程度，参数 n_p 越多，越分散，方差越大

$$E(\hat{y}_1) - y$$

表示估计值和真实值的差，参数越多，模型越准确，偏差越小

$$MSE(z - \tilde{y}_1)$$

$$MSE(z - \tilde{y}_1) = Var(z - \tilde{y}_1) + [E(\tilde{y}_1) - y]^2$$



随着参数量增加，方差增大，偏差误差减小。存在一个合适值，使得均方误差最小



假设检验：分析超过一个参数的统计模型，以判断模型中全部或部分参数是否合适来估计母体。

基本思想：先提出假设检验，再用适当的统计学方法，利用小概率原理确定假设是否成立。即为了检验一个假设 H_0 是否正确，先假定 H_0 正确，然后根据样本对假设 H_0 做出接收或拒绝的决策。如果样本量测值导致 H_0 为一个小概率事件，则应该拒绝 H_0 。

主要步骤：

1. 提出假设检验，又称无效检验 H_0 以及备设检验 H_1

$$H_0: \theta_1 = \theta_2 = \cdots = \theta_n = 0$$

$$H_1: \text{至少存在一个参数 } \theta_j \neq 0$$

如果 H_0 被拒绝，代表只要有一个回归量 ξ_j 对描述模型贡献显著

2. 预先设定的检验水平为 α ，表示假设 H_0 为真的概率为 α ，通常 $\alpha = 0.05$ ；
3. 选择合适的检验方法，由量测值 z 按照相应的公式计算出统计量概率的大小；
4. 若 $P > \alpha$ ，则不拒绝 H_0 ，若 $P \leq \alpha$ ，则拒绝 H_0 ，接收 H_1



F检验（联合假设检验）：用来分析超过一个参数的统计模型，判断该模型中的全部或一部分参数是否合适用来估计母体。

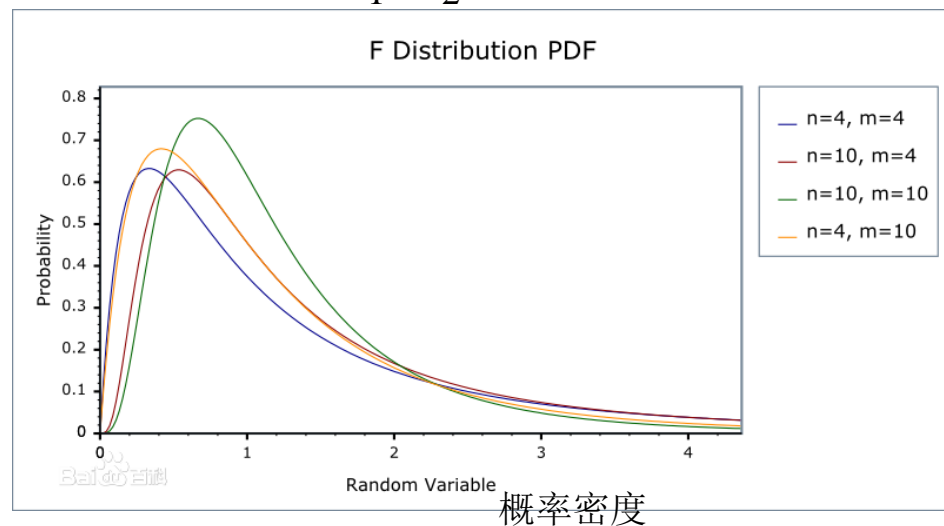
思路：用于分析在相同母体中取出两组数据的差异性，判断两组数据的方差是否存在显著差异。



设总体 X 服从 $X \sim N(0,1)$, $(X_1, X_2, \dots, X_{n_1})$ 和 $(Y_1, Y_2, \dots, Y_{n_2})$ 为来自 X 的两个独立样本, 则

$$F = \frac{\sum_{i=1}^{n_1} X_i^2}{n_1} \bigg/ \frac{\sum_{i=1}^{n_2} Y_i^2}{n_2}$$

称统计量 F 服从自由度 n_1 和 n_2 的 F 分布, 记为 $F \sim F(n_1, n_2)$





回归量显著性检测

1. 提出无效检验 H_0 以及备设检验 H_1

$$H_0: \theta_1 = \theta_2 = \dots = \theta_n = 0$$

$$H_1: \text{至少存在一个参数 } \theta_j \neq 0$$

2. 计算 F_0

模型中仅包含 θ_0 时的 $\hat{\sigma}^2$

$$F_0 = \frac{\frac{SS_R/n}{\frac{SS_E/(N - n_p)}{\hat{\sigma}^2}}}{\hat{\sigma}^2} = \frac{\hat{\theta}^T X^T Z - N \bar{Z}^2}{ns^2}$$

3. 若 $F_0 > F(\alpha, n, N - n_p)$, 拒绝 H_0 , 表示 H_1 有 $(1 - \alpha)$ 的置信度



部分参数显著性检测

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{bmatrix} \quad \boldsymbol{\theta}_1 \text{ 为 } p \times 1 \text{ 矩阵, 包括 } \theta_0; \quad \boldsymbol{\theta}_2 \text{ 为 } q \times 1 \text{ 矩阵}$$

$$p + q = n_p$$

$$\mathbf{z} = \mathbf{X}_1 \boldsymbol{\theta}_1 + \mathbf{X}_2 \boldsymbol{\theta}_2 + \mathbf{v}$$

1. 提出无效检验 H_0 以及备设检验 H_1

$$H_0: \boldsymbol{\theta}_2 = \mathbf{0}$$

$$H_1: \boldsymbol{\theta}_2 \neq \mathbf{0}$$

2. 计算 F_0

对于全参数模型, 已知

$$\hat{\boldsymbol{\theta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{z}$$

$$SS_R = \hat{\boldsymbol{\theta}}^T \mathbf{X}^T \mathbf{z} - N \bar{z}^2$$

$$SS_E = \mathbf{z}^T \mathbf{z} - \hat{\boldsymbol{\theta}}^T \mathbf{X}^T \mathbf{z}$$



部分参数显著性检测

2. 计算 F_0

为评价 $X_2\theta_2$ 对模型的贡献，需要假设 H_0 成立的前提下，估计模型参数 $\hat{\theta}_1$ 。

$$z = X_1\theta_1 + v$$

$$\hat{\theta}_1 = (X_1^T X_1)^{-1} X_1^T z$$

$$SS_R(\hat{\theta}_1) = \hat{\theta}_1^T X_1^T z - N \bar{z}^2$$

$$SS_E(\hat{\theta}_1) = z^T z - \hat{\theta}_1^T X_1^T z$$

在 $X_1\theta_1$ 已经存在的前提下，由 $X_2\theta_2$ 产生的 $SS_R(\hat{\theta}_2|\hat{\theta}_1)$

$$SS_R(\hat{\theta}_2|\hat{\theta}_1) = SS_R - SS_R(\hat{\theta}_1) = \hat{\theta}^T X^T z - \hat{\theta}_1^T X_1^T z$$

$$F_0 = \frac{SS_R(\hat{\theta}_2|\hat{\theta}_1)}{qs^2}$$

3. 若 $F_0 > F(\alpha, n, N - n_p)$ ，拒绝 H_0 ，表示 H_1 有 $(1 - \alpha)$ 的置信度



单一参数显著性检测

1. 提出无效检验 H_0 以及备设检验 H_1

$$H_0: \theta_2 = 0$$

$$H_1: \theta_2 \neq 0$$

2. 计算 F_0

$$\text{定义 } z_1 = z - X_1 \hat{\theta}_1 = X_2 \theta_2 + v$$

$$F_0 = \frac{\hat{\theta}_2 X_1^T z_1}{s^2} = \frac{\hat{\theta}_2^2}{s^2(\hat{\theta}_2)}$$

3. 若 $F_0 > F(\alpha, n, N - n_p)$, 拒绝 H_0 , 表示 θ_2 有 $(1 - \alpha)$ 的置信度

三、模型选择



前向筛选

该技术始于没有回归变量，只有常数项 θ_0 。然后一次一个回归直到所有候选回归变量都在模型中或满足某些选择条件为止。

$$F_0 = \frac{SS_R(\hat{\theta}_{p+j}) - SS_R(\hat{\theta}_p)}{s^2} = \frac{SS_R(\hat{\theta}_j | \hat{\theta}_p)}{s^2} \geq F_{in}$$

满足上式，将参数加入模型

其中： SS_R 是回归平方和， s^2 是将参数 j 考虑在内的拟合误差，用来表示估计值接近真实值的程度。

F_{in} 是 $F(\alpha; 1, N - p - 1)$

$$SS_R = \sum_{i=1}^N [\hat{y}(i) - \bar{z}]^2$$

$$s^2 = \hat{\sigma}^2 = \frac{\mathcal{U}\mathcal{U}^T}{(N - n_p)}$$

三、模型选择



前向筛选

注意：优先筛选与量测值线性相关性最高回归量。

回归量的方差 $S_{jj} = \sum_{i=1}^N [\xi_j(i) - \bar{\xi}_j]^2$

量测量的方差 $S_{zz} = \sum_{i=1}^N [z(i) - \bar{z}]^2$

不同回归量之间的相关性系数

$$r_{jk} = \frac{\sum_{i=1}^N [\xi_j(i) - \bar{\xi}_j][\xi_k(i) - \bar{\xi}_k]}{\sqrt{S_{jj}S_{kk}}}$$

回归量与量测量相关性系数

$$r_{jz} = \frac{\sum_{i=1}^N [\xi_j(i) - \bar{\xi}_j][z(i) - \bar{z}]}{\sqrt{S_{jj}S_{zz}}}$$



后向删除

该技术始于回归方程中包含所有的候选回归变量。然后，不必要的参数一次删除一个

$$F_0 = \min_j \frac{SS_R(\hat{\theta}_p) - SS_R(\hat{\theta}_{p-j})}{s^2} \leq F_{out}$$

满足上式，将该参数去除

其中， F_{out} 是 $F(\alpha; 1, N - p - 1)$

注意：优先删除与量测值线性相关性最小的回归量。



逐步回归：前向筛选和后向删除的结合

1. 初始状态：仅包含偏差项 θ_0 。
2. 选择第一个参数 ξ_1 后，模型为

$$z(i) = \theta_0 + \theta_1 \xi_1(i) + \nu(i) \quad i=1,2,3,\dots,N$$

$$F_0 = \frac{SS_R(\theta_1)}{S^2}$$

如果 F_0 大于 F_{in} ，那么将 ξ_1 引入模型



逐步回归：前向筛选和后向删除的结合

2. 接着添加第二个参数 ξ_2 时，我们需消除第一个参数 ξ_1 对第二个参数 ξ_2 的影响

定义第一个参数 ξ_1 与第二个参数 ξ_2 关系为：

$$\xi_2(i) = a_0 + a_1 \xi_1(i) + v_2(i)$$

那么定义新的回归量 v_2 （ ξ_2 去除了常数以及与 ξ_1 的线性相关部分）

$$v_2(i) = \xi_2(i) - \hat{a}_0 - \hat{a}_1 \xi_1(i)$$

其中， \hat{a}_0 、 \hat{a}_1 为最小二乘估计结果

同理，

$$z_2(i) = \mathbf{z}(i) - \hat{\theta}_0 - \hat{\theta}_1 \xi_1(i)$$

单一参数显著性检测

$$F_0 = \frac{\hat{\theta}_2 \mathbf{X}_1^T \mathbf{z}_1}{s^2} = \frac{\hat{\theta}_2^2}{s^2(\hat{\theta}_2)}$$

如果 F_0 大于 F_{in} ，那么将 ξ_2 引入模型



逐步回归：前向筛选和后向删除的结合

3. 每引入一个新的变量 ξ_j 进入模型，将已包含在模型里面的回归量和新回归量带入下式。

$$z(i) = \theta_0 + \sum_{j=1}^n \theta_j \xi_j(i) + v(i) \quad i = 1, 2, \dots, N$$

4. 后向删除，去除影响很小的或者与其他项过分线性相关的参数

$$F_0 = \min_j \frac{SS_R(\hat{\theta}_p) - SS_R(\hat{\theta}_{p-j})}{s^2} \leq F_{out}$$

利用上式逐个对参数进行评估，满足上式的参数，去除
删除参数的过程一直持续到没有更多的参数

一般情况下： $F_{in} > F_{out}$ ，增加一个回归量比去掉回归量要更难。



其他停止模型筛选的准则

根据实测数据仅将部分F统计量与Fin和Fout一起使用有时过于严格。
这里介绍几个其他的选择模型结构的准则：

1.确定系数 R^2

$$R^2 = \frac{SS_R}{SS_T} = \frac{\hat{\theta}^T X^T z - N\bar{z}^2}{z^T z - N\bar{z}^2}$$

增加参数， R^2 增大，如果参数的影响越大， R^2 增大的幅度越大。如果增加一个参数 R^2 增大的幅度小于0.5%，那么此参数影响很小，可以忽略

总平方和 $SS_T = \sum_{i=1}^N [z(i) - \bar{z}]^2$

回归平方和 $SS_R = \sum_{i=1}^N [\hat{y}(i) - \bar{z}]^2$

残差平方和 $SS_E = \sum_{i=1}^N [z(i) - \hat{y}(i)]^2$

有 $SS_R + SS_E = SS_T$



其他停止模型筛选的准则

2.F值

$$F0 = \frac{SS_R / n}{SS_E / (N - p)} = \frac{\hat{\theta}^T X^T z - N \bar{z}^2}{ns^2}$$

其中，n是删减此参数时的参数数量，n=p-1

F0与R²的关系为：

$$F0 = \frac{(N - p)}{(p - 1)} \frac{R^2}{1 - R^2}$$

这种关系表明模型开始时F0的变化选择过程可以非常明显。但是后来，当影响力降低时项添加到模型中，R²/(1-R²)的增加可能小于(N-p)(p-1)的减小。结果，F0的值存在一个最大值。**F0=F0_{max}可以作为一个终止准则。**



其他停止模型筛选的准则

4. 预测平方误差, PSE

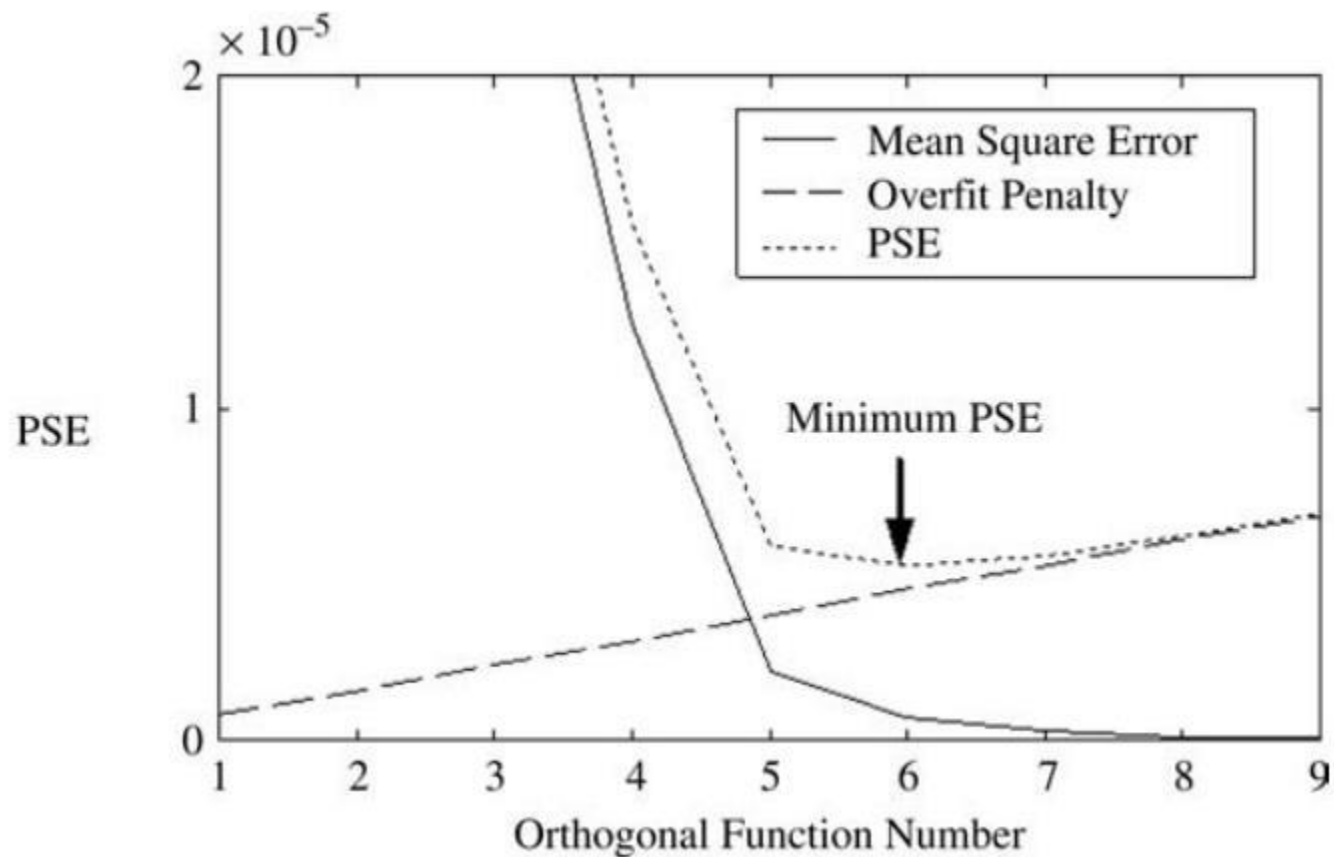
$$\begin{aligned} PSE &= \frac{1}{N} (z - \hat{y})^T (z - \hat{y}) + \sigma_{\max}^2 \frac{p}{N} \\ &= MSFE + \sigma_{\max}^2 \frac{p}{N} \end{aligned}$$

MSFE是均方拟合误差, $\sigma_{\max}^2 \frac{p}{N}$ 是超额惩罚。P为模型中回归量个数, 增加参数后, MSFE减小, 超额惩罚增大。

终止准则为: $PSE = \min PSE$

其中, σ_{\max}^2 为估计方差的一个上界,

$$\sigma_{\max}^2 = \frac{1}{N} \sum_{i=1}^N [z(i) - \bar{z}]^2$$

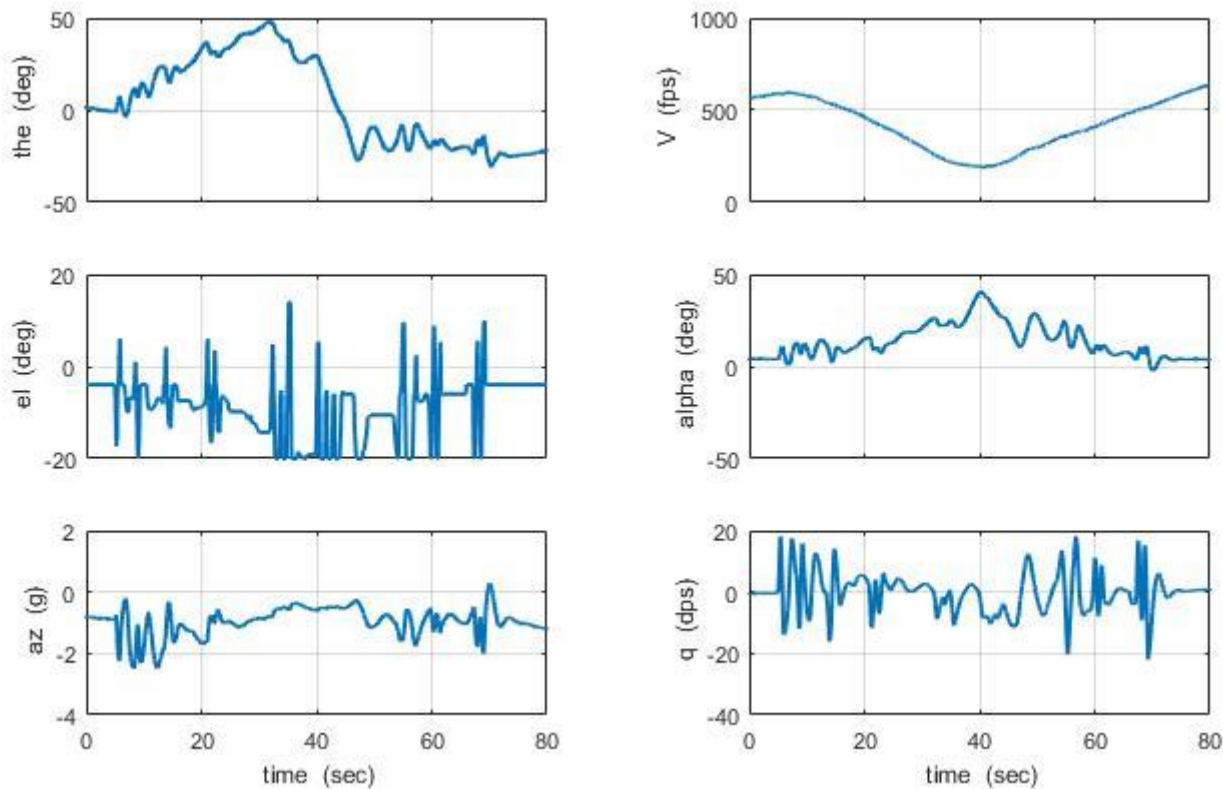


PSE值随参数量的变化



例子:

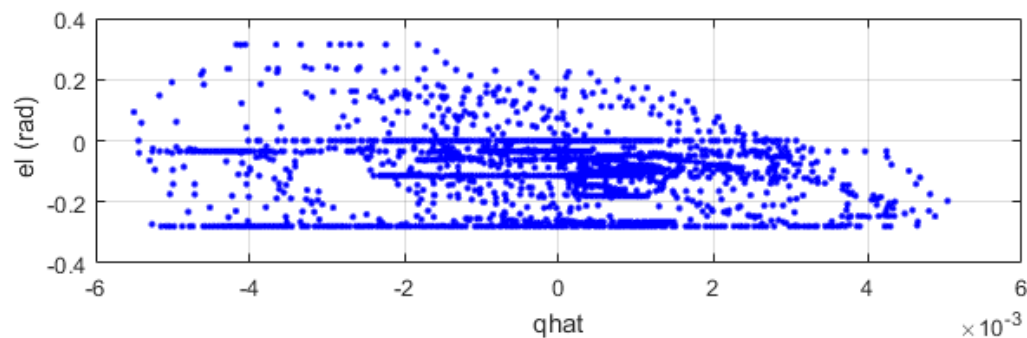
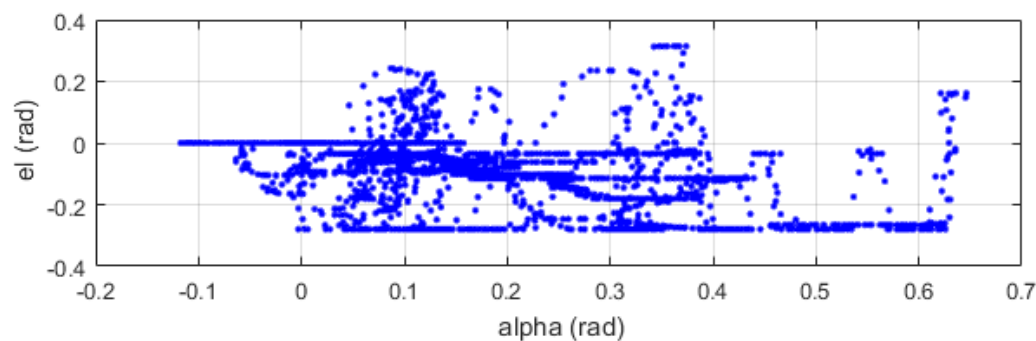
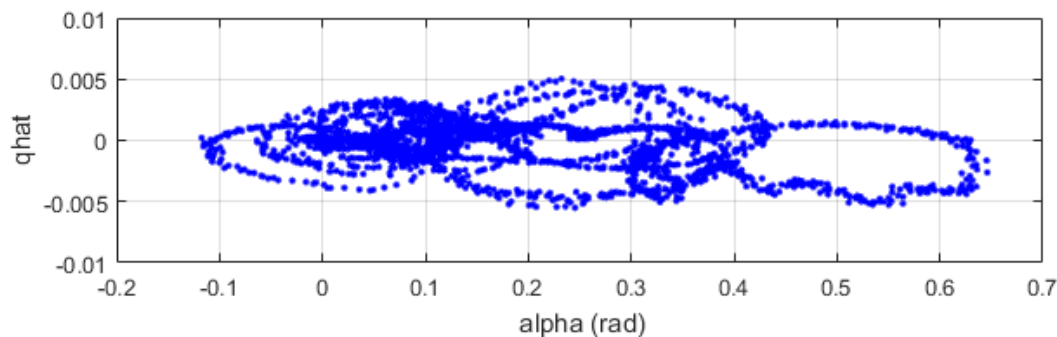
大幅度纵向机动下的输入输出为



三、模型选择

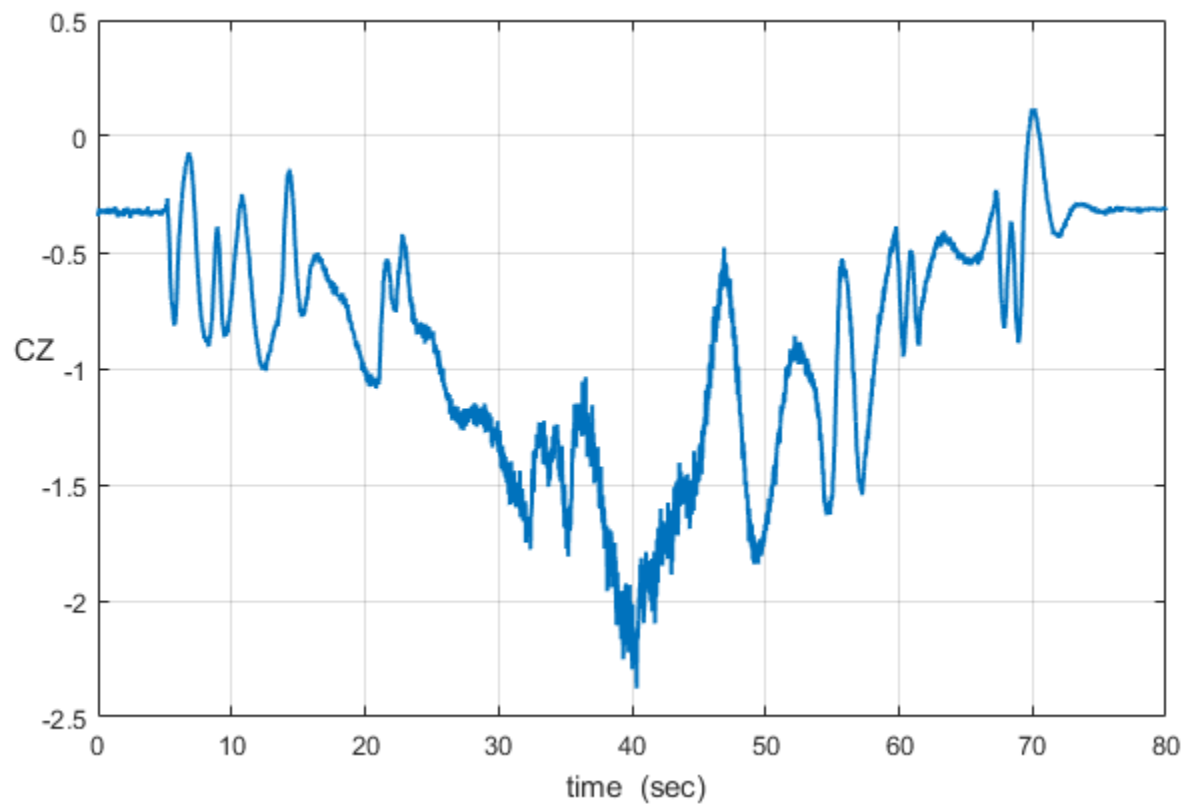


考察输入输出间的
相关性：若为
直线，说明不相
关





真实输出偏航
力矩系数 C_z



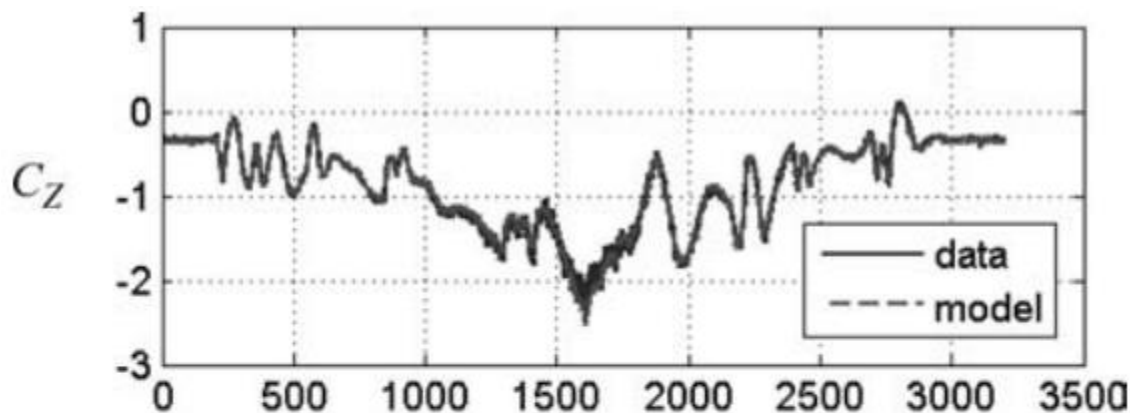
三、模型选择



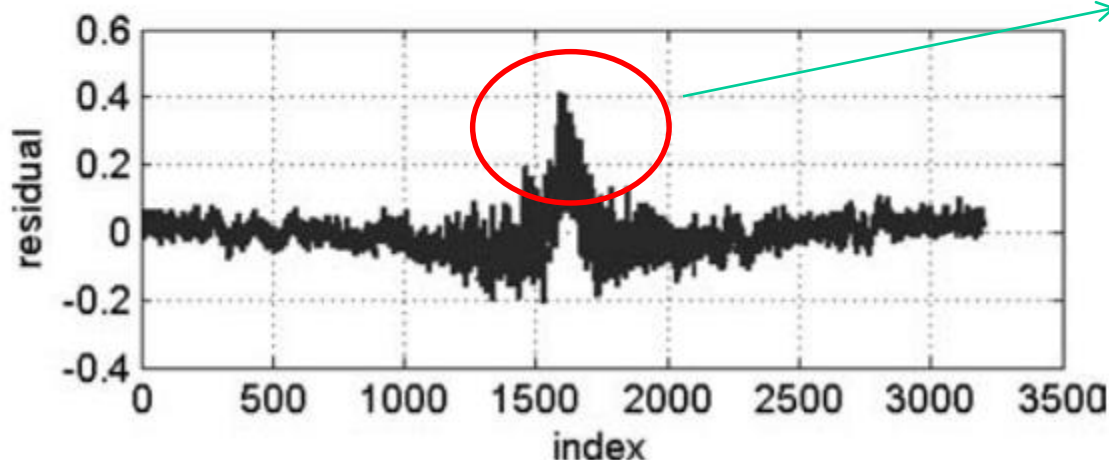
逐步回归的结果如下：

Step	1			2			3			4		
	$\hat{\theta}$	F_0	r	$\hat{\theta}$	F_0	r	$\hat{\theta}$	F_0	r	$\hat{\theta}$	F_0	r
C_{Z_M}	0	—	0.725	0	—	0.020	0	—	0.000	0	—	0.100
C_{Z_α}	0	—	0.973	−3.13	1.2e5	—	−3.21	1.6e5	—	−3.35	2.0e5	—
C_{Z_q}	0	—	0.031	0	—	0.305	−27.0	1.4e3	—	−32.3	2.8e3	—
$C_{Z_{\delta_e}}$	0	—	0.166	0	—	0.114	0	—	0.322	−0.42	1.5e3	—
$C_{Z_{\alpha^2}}$	0	—	0.838	0	—	0.306	0	—	0.303	0	—	0.485
PRESS	852.0			23.0			16.0			10.9		
PSE	0.266			0.00735			0.00524			0.00372		
$R^2, \%$	0.0			97.3			98.1			98.7		
5												

上述第四步得到的结果如下：



模型拟合地很好

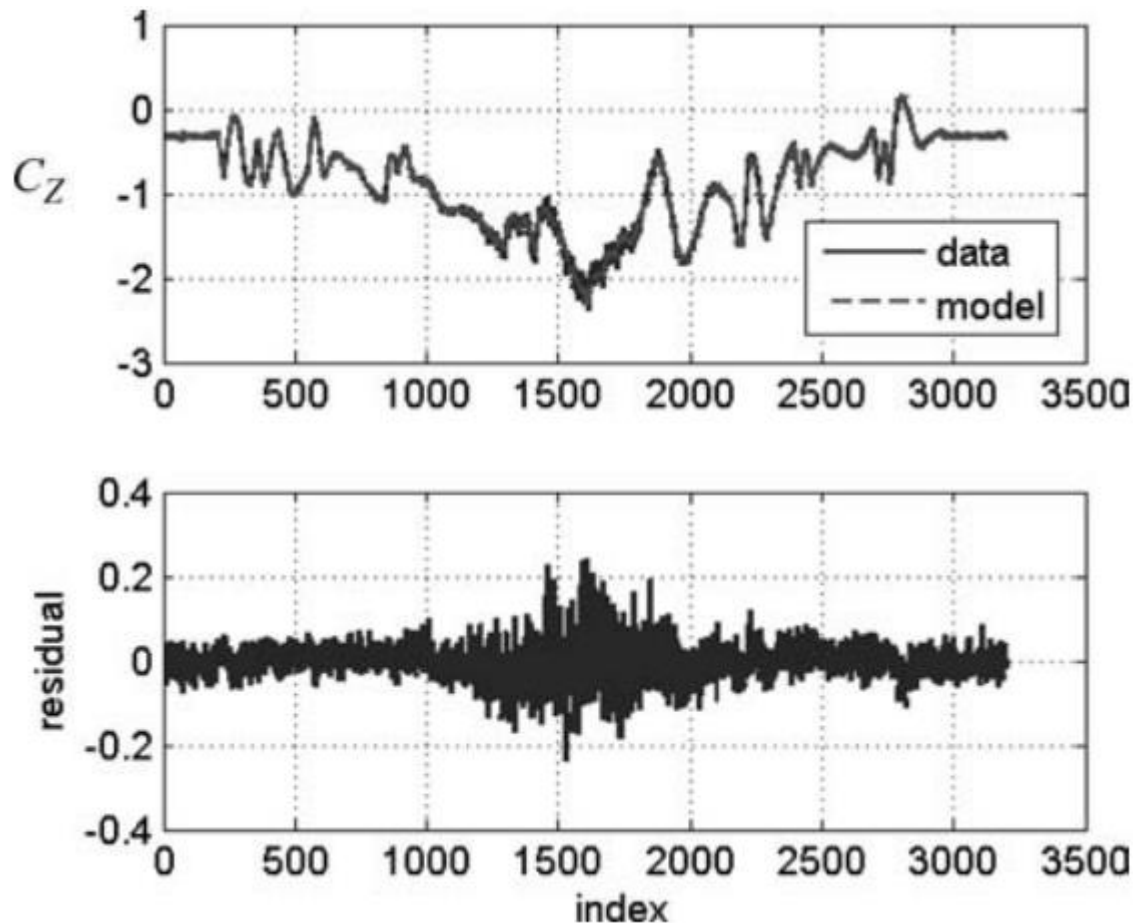


残差不符合白噪声假设

$$C_Z = C_{Z_M} M + C_{Z_\alpha} \alpha + C_{Z_q} q + C_{Z_{\delta_e}} \delta_e$$



第五步，再加入参数 α^2 后



$$C_z = C_{z_M} M + C_{z_\alpha} \alpha + C_{z_q} q + C_{z_{\delta_e}} \delta_e + C_{z_{\alpha^2}} \alpha^2$$

三、模型选择



Parameter	$\hat{\theta}$	$s(\hat{\theta})$	$ t_0 $	$100[s(\hat{\theta})/ \hat{\theta}]$
C_{Z_M}	0	—	—	—
C_{Z_α}	-4.166	0.0497	83.9	1.2
C_{Z_q}	-28.13	0.9822	28.6	3.5
$C_{Z_{\delta_e}}$	-0.4333	0.0123	35.3	2.8
$C_{Z_{\alpha^2}}$	1.344	0.0823	16.3	6.1
C_{Z_o}	-0.3229	0.0034	95.0	1.1
$s = \hat{\sigma}$	0.0418	—	—	—
$R^2, \%$	99.3	—	—	—

第四章：飞行器时域辨识方法

1. 基础知识
2. 最小二乘方法
3. 基于最小二乘方法的模型选择
4. 最大似然方法
5. 时域辨识试验设计

四、最大似然估计方法——基本知识




在飞行器系统识别的范围内，有一套完善的最大似然法算法

→ 基础是Fisher模型的最优估计

设 $p(z_k | \theta)$ 为量测 z_k 的概率密度，其中包含一系列确定性的参数 $\theta = [\theta_1, \theta_2, \dots, \theta_{n_\theta}]^T$ 。通过调节参数 θ ，使得在 t_k 时刻，获得量测 $z = [z_1, z_2, \dots, z_N]$ 的概率最大。

$$\max_{\theta} p(Z | \theta)$$

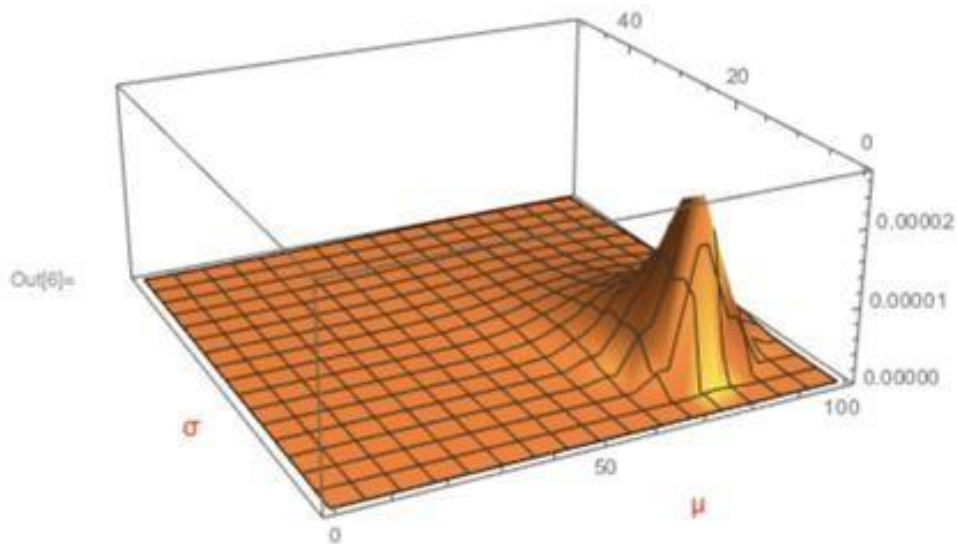

$$\hat{\theta} = \hat{\theta}(z) = \arg \max_{\theta} p(Z | \theta)$$

在每个时刻 k ， $k=1, 2, \dots, N$ ，都有 n_z 个量测已知：

四、最大似然估计方法——基本知识



通过调整参数（峰所在位置），直到得到测量最大的可能性





四、最大似然估计方法——基本知识

在统计学中，在 x_1, x_2 独立的时候：

$$p(x_1, x_2) = p(x_1)p(x_2)$$

假设：测量 z_k 是独立的且具有相同分布

1. 他们是统计学独立的（在时间上不相关/白噪声）

→ k 时刻的测量的随机部分与 $k+1$ 时刻的测量的随机部分是独立的

2. 他们具有相同分布

→ $p(z_k | \theta) = p(z_j | \theta)$ 对于所有 $k, j = 1, \dots, N$

所以 $p([z_k, z_{k+1}] | \theta) = p(z_k | \theta) \cdot p(z_{k+1} | \theta)$

因此，测量 Z 可以表示为：

$$p(Z | \theta) = \prod_{k=1}^N p(z_k | \theta)$$



四、最大似然估计方法——基本知识

现在，残差 r 可以假设为：

1. 0均值：

$$E(r)=0$$

2. 高斯分布且协方差矩阵为 R ：

$$\text{Cov}[r]=R$$

他们与测量 z_k 以及模型输出 $h_k(\theta, u, \tilde{w})$ 有关

$$r_k = z_k - h_k(\theta, u)$$

因此协方差矩阵为：

附加的噪声情况 $R = \text{Cov}[z - h(\theta, u) - v] = \text{Cov}[v]$

在 t_k 时刻获得测量 z_k 的概率完全取决于 $r_k = z_k - h_k(\theta, u)$ 和多变量高斯分布的协方差矩阵 R

$$p(z_k | \theta) = p(r_k) = \frac{1}{\sqrt{(2\pi)^{n_z} |B|}} \exp\left(-\frac{1}{2} \underbrace{(z_k - h_k(\theta, u))^T}_{r_k^T} R^{-1} \underbrace{(z_k - h_k(\theta, u))}_{r_k}\right)$$



四、最大似然估计方法——基本知识

因此，似然函数 $L(z;\theta)$ 可以表示为：

$$\begin{aligned} L(z;\theta) &= p(z|\theta) = \prod_{k=1}^N p(z_k|\theta) \\ &= ((2\pi)^{n_z} |R|)^{-\frac{N}{2}} \exp\left(-\frac{1}{2} \sum_{k=1}^N (z_k - h_k(\theta, u, \tilde{w}))^T R^{-1} (z_k - h_k(\theta, u))\right) \end{aligned}$$

在进行最优求解的过程中，偏导数将被用到，求导规则如下：

$$\begin{aligned} \frac{\partial L(z;\theta)}{\partial \theta} &= \frac{\partial p(z_1|\theta)}{\partial \theta} \prod_{k=2}^N p(z_k|\theta) + \dots + \frac{\partial p(z_i|\theta)}{\partial \theta} \prod_{k=1, k \neq i}^N p(z_k|\theta) + \dots + \frac{\partial p(z_N|\theta)}{\partial \theta} \prod_{k=1}^{N-1} p(z_k|\theta) \\ &= \sum_{i=1}^{n_\theta} \frac{\partial p(z_i|\theta)}{\partial \theta} \prod_{k=1, k \neq i}^N p(z_k|\theta) \end{aligned}$$



四、最大似然估计方法——基本知识

通过取对数简化这个问题，对似然函数取对数得：

$$-\ln(L(z; \theta)) = -\ln\left(\prod_{k=1}^N p(z_k | \theta)\right) = -\sum_{i=1}^N \ln(p(z_k | \theta))$$

再求偏导数：

$$\frac{\partial(-\ln(L(z; \theta)))}{\partial \theta} = -\frac{\partial}{\partial \theta} \left(\sum_{i=1}^N \ln(p(z_k | \theta)) \right) = -\sum_{i=1}^N \frac{\partial \ln(p(z_k | \theta))}{\partial \theta}$$

似然函数中的指数在取对数求导的过程中被消去，得到更为简单的结果如下：

$$\ln p(z_k | \theta) = -\frac{1}{2} \ln((2\pi)^{n_z} | R |) - \frac{1}{2} (z_k - h_k(\theta, u))^T R^{-1} (z_k - h_k(\theta, u))$$

四、最大似然估计方法——基本知识



转换 $L(z;\theta)$

代价函数 $L(z;\theta)$ 的转换是有效的，由于：

1). 求一个函数的最大值等效于求其负对数的最小值

(绝大多数最优问题被转换成求最小值问题)

2). 自然对数是一个严格单调的函数。求对数形式的代价函数的最大值等同于求原函数的最大值



四、最大似然估计方法——基本知识

转换 $L(z;\theta)$

根据前面介绍的，自然函数求最大值转换成如下形式：

$$\begin{aligned}
 -\ln(L(z;\theta)) &= -\ln\left(\prod_{k=1}^N p(z_k | \theta)\right) \\
 &= -\ln\left(\exp\left(-\frac{1}{2} \sum_{i=1}^N (z_k - h_k(\theta, u, \tilde{w}))^T B^{-1} (z_k - h_k(\theta, u, \tilde{w}))\right) \cdot ((2\pi)^{n_z} |B|)^{-\frac{N}{2}}\right) \\
 &= \frac{1}{2} \sum_{i=1}^N (z_k - h_k(\theta, u, \tilde{w}))^T B^{-1} (z_k - h_k(\theta, u, \tilde{w})) - \ln((2\pi)^{n_z} |B|)^{-\frac{N}{2}}
 \end{aligned}$$

最终结果为：

$$\min_{\theta} J(\theta) \longleftrightarrow \hat{\theta} = \arg \max_{\theta} J(\theta)$$

$$J(\theta) = -\ln(L(z;\theta))$$

其中：

$$= \frac{1}{2} \sum_{i=1}^N (z_k - h_k(\theta, u, \tilde{w}))^T B^{-1} (z_k - h_k(\theta, u, \tilde{w})) + \frac{N}{2} \ln(|B|) + \frac{Nn_z}{2} \ln(2\pi)$$

四、最大似然估计方法——基本知识



Fisher模型-总结

Fisher模型:

1. θ 是未知常值参数的向量
2. r 是随机向量, 其概率密度为 $p(r)$

1. 估计基于似然函数 $L(z; \theta) = p(z|\theta)$

2. 当 r 符合高斯分布时, 似然函数为:

$$L(z; \theta) = \left((2\pi)^{n_z} |R| \right)^{-\frac{N}{2}} \exp \left(-\frac{1}{2} \sum_{k=1}^N (z_k - h_k(\theta, u))^T R^{-1} (z_k - h_k(\theta, u)) \right)$$

最大似然估计 θ_{ML} 使代价函数最小求解得到

$$\hat{\theta}_{ML} = \min_{\theta} J(\theta)$$

$$J(\theta) = \frac{N}{2} \ln(|R|) + \frac{N n_z}{2} \ln(2\pi) + \frac{1}{2} \sum_{i=1}^N (z_k - h_k(\theta, u))^T R^{-1} (z_k - h_k(\theta, u))$$



四、最大似然估计方法——基本知识

最大似然到底干了什么？

1.在试验中，用一个模型去拟合量测值



2.过程中的随机部分（过程噪声和测量噪声）



3.基本思想：调整模型中的参数，使得得到测量值的概率最大



4.假设样本独立使得事件发生的概率为各个样本发生概率的乘积



5.假设残差 r 符合高斯分布，以及其给定的协方差矩阵使得结果进一步简化



6.在求解代价函数问题中使用负对数，使得计算更加简单



四、最大似然估计方法——基本知识

最大似然估计

非线性优化：

$$\min_x f(x)$$

求解方法：局部最小值

梯度

$$\left. \frac{df(x)}{dx} \right|_{x=\hat{x}_i} = 0$$

Hessian矩阵

$$H(x) = \left. \frac{d^2 f(x)}{dx^2} \right|_{x=\hat{x}_i} = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n \partial x_n} \end{bmatrix} > 0$$

四、最大似然估计方法——基本知识

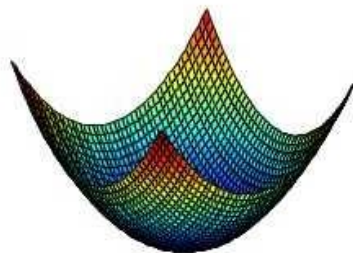
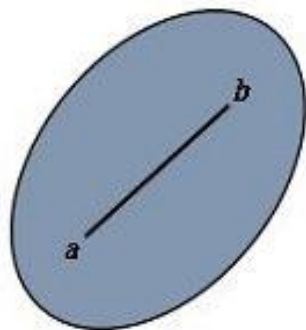
凸优化

定义在某个向量空间的凸子集上的实值函数 f ，如果其定义域上的任意两点 x, y ，以及 $t \in (0, 1)$ ，有

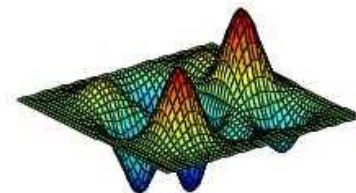
$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y)$$

如果目标函数 J 是一个凸函数，则这样的优化问题为凸优化。

对于凸优化，局部最优解就是全局最优解。



- Unique optimum: global/local.



- Multiple local optima
- In high dimensions possibly exponential local optima

四、最大似然估计方法——基本知识

梯度下降法

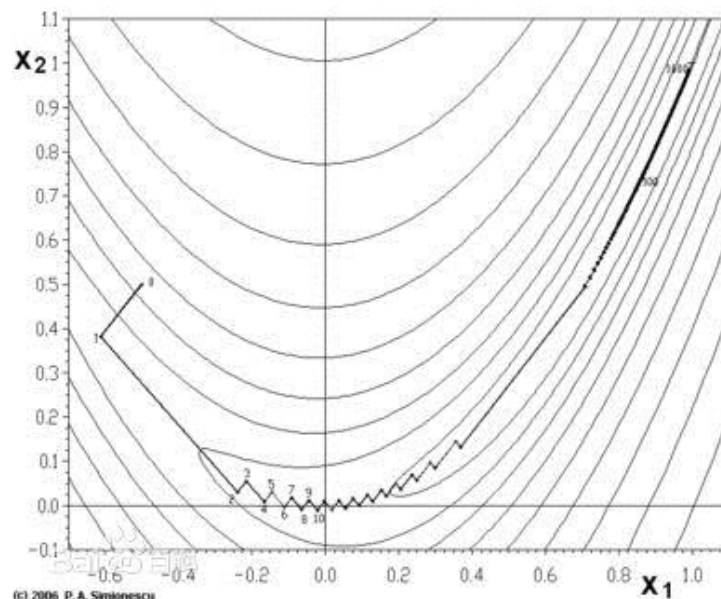
梯度：对于可微的函数 $f(x, y, z)$ ，则 $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$ 为分量的向量称为 $f(x, y, z)$ 的梯度。

梯度下降法：沿梯度下降的方向求解极小值

$$f_{k+1} = f_k + a_k \bar{s}(k)$$

$\bar{s}(k)$ ：梯度的负方向， a_k 梯度方向上的搜索步长。

步长太大可能发散，太小收敛速度太慢。





四、最大似然估计方法——基本知识

求解最大似然问题

代价函数：

$$J(\boldsymbol{\theta}) = \frac{1}{2} \sum_{k=1}^N (z_k - h_k(\theta))^T R^{-1} (z_k - h_k(\theta)) + \frac{N}{2} \ln(|\mathbf{R}|) + \frac{Nm}{2} \ln(2\pi)$$

问题：

如何求解非线性代价函数？ → **Newton-Raphson**算法



四、最大似然估计方法——Newton-Raphson算法

问题：必要条件 梯度 $\left. \frac{df(x)}{dx} \right|_{x=\hat{x}_i} = 0$ 通常不能通过分析得到。

Newton-Raphson算法是将梯度基于一阶泰勒公式展开：

$$\left. \frac{\partial f(x)}{\partial x} \right|_{x=x_{i+1}} = \left. \frac{\partial f(x)}{\partial x} \right|_{x=x_i} + \left. \frac{\partial^2 f(x)}{\partial x^2} \right|_{x=x_i} \boxed{(x_{i+1} - x_i)} = 0$$

因此，变量可根据此式由初始值 x_0 迭代计算：

$$\Delta x_i = - \left(\left. \frac{\partial^2 f(x)}{\partial x^2} \right|_{x=x_i} \right)^{-1} \left. \frac{\partial f(x)}{\partial x} \right|_{x=x_i} \quad x_{i+1} = x_i + \Delta x_i$$

通常满足收敛判据时， $\frac{f(x_{i+1}) - f(x_i)}{f(x_i)} \leq f_{Conv}$ 或 $\frac{\Delta x_i}{x_i} \leq \Delta x_{conv}$
迭代停止

四、最大似然估计方法——Newton-Raphson算法



最大似然代价函数

$$J(\boldsymbol{\theta}) = \frac{1}{2} \sum_{i=1}^N \boldsymbol{v}^T(i) \boldsymbol{R}^{-1} \boldsymbol{v}(i) + \frac{N}{2} |\boldsymbol{R}|$$

二阶泰勒展开

$$J(\boldsymbol{\theta}_0 + \Delta\boldsymbol{\theta}) = J(\boldsymbol{\theta}_0) + \Delta\boldsymbol{\theta}^T \left. \frac{\partial J}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}_0} + \Delta\boldsymbol{\theta}^T \left. \frac{\partial^2 J}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \Delta\boldsymbol{\theta} + \dots$$

$\frac{\partial^2 J}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}$ 为 Hessian 矩阵

二阶截断

$$J(\boldsymbol{\theta}_0 + \Delta\boldsymbol{\theta}) \approx J(\boldsymbol{\theta}_0) + \Delta\boldsymbol{\theta}^T \left. \frac{\partial J}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}_0} + \Delta\boldsymbol{\theta}^T \left. \frac{\partial^2 J}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \Delta\boldsymbol{\theta}$$

四、最大似然估计方法——Newton-Raphson算法



$J(\theta_0 + \Delta\theta)$ 取极小的必要条件

$$\frac{\partial}{\partial \theta} J(\theta_0 + \Delta\theta) = 0$$

即：

$$\frac{\partial}{\partial \theta} J(\theta_0 + \Delta\theta) = \left. \frac{\partial J}{\partial \theta} \right|_{\theta_0} + \left. \frac{\partial^2 J}{\partial \theta \partial \theta^T} \right|_{\theta=\theta_0} \Delta\theta = 0$$

$$\Delta\hat{\theta} = - \left[\left. \frac{\partial^2 J}{\partial \theta \partial \theta^T} \right|_{\theta=\theta_0} \right]^{-1} \left[\left. \frac{\partial J}{\partial \theta} \right|_{\theta_0} \right]$$

θ_0 的参数更新值 $\hat{\theta}$ 为 $\hat{\theta} = \theta_0 + \Delta\hat{\theta}$

问题：如何求解目标函数J梯度和Hessian矩阵



四、最大似然估计方法——Newton-Raphson算法

残差R的协方差矩阵

目标函数

$$J(\boldsymbol{\theta}) = \frac{1}{2} \sum_{i=1}^N \mathbf{v}^T(i) \mathbf{R}^{-1} \mathbf{v}(i) + \frac{N}{2} |\mathbf{R}|$$

理想化:

- R只考虑在使用中传感器量测噪声的影响。

实际上:

- R还包括模型输出 \mathbf{y}_k 和测量值 \mathbf{z}_k 之间的差（目前的参数还没涉及到这两方面）；
- R还包括确定性模型误差（任何模型都不能精准地描述真实情况）。

R计算为模型输出函数
→取决于参数 $\boldsymbol{\theta}$

$\boldsymbol{\theta}$ 通过代价函数计算
→取决于R

实际上：（松弛方法）

通过固定值 $\boldsymbol{\theta}_i$ 计算 \mathbf{R}_i ；

通过固定值 \mathbf{R}_i 确定参数变化 $\Delta \boldsymbol{\theta}_i$ 。

$$\mathbf{R} = \frac{1}{N} \sum_{k=1}^N \mathbf{v}_k \mathbf{v}_k^T = \frac{1}{N} \sum_{k=1}^N (\mathbf{z}_k - \mathbf{y}_k)(\mathbf{z}_k - \mathbf{y}_k)^T$$

四、最大似然估计方法——Newton-Raphson算法



求解最大似然问题

目标函数

$$J(\boldsymbol{\theta}) = \frac{1}{2} \sum_{i=1}^N \mathbf{v}^T(i) \hat{\mathbf{R}}^{-1} \mathbf{v}(i) + \frac{N}{2} |\hat{\mathbf{R}}| = \frac{1}{2} \sum_{i=1}^N [\mathbf{z}(i) - \mathbf{y}(i)]^T(i) \hat{\mathbf{R}}^{-1} [\mathbf{z}(i) - \mathbf{y}(i)]$$

与未知参数 $\boldsymbol{\theta}$ 无关

梯度

$$\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = - \sum_{i=1}^N \frac{\partial \mathbf{y}^T(i)}{\partial \boldsymbol{\theta}} \hat{\mathbf{R}}^{-1} \mathbf{v}(i)$$

梯度矩阵中的第j个元素为

$$\frac{\partial J(\boldsymbol{\theta})}{\partial \theta_j} = - \sum_{i=1}^N \frac{\partial \mathbf{y}^T(i)}{\partial \theta_j} \hat{\mathbf{R}}^{-1} \mathbf{v}(i) \quad j = 1, 2, \dots, n_p$$

Hessian 矩阵的元素为

$$\frac{\partial^2 J(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_k} = \sum_{i=1}^N \frac{\partial \mathbf{y}^T(i)}{\partial \theta_j} \hat{\mathbf{R}}^{-1} \frac{\partial \mathbf{y}(i)}{\partial \theta_k} - \sum_{i=1}^N \frac{\partial^2 \mathbf{y}(i)}{\partial \theta_j \partial \theta_k} \hat{\mathbf{R}}^{-1} \mathbf{v}(i)$$

$$j, k = 1, 2, \dots, n_p$$

1. 二阶导计算困难
2. $\hat{\boldsymbol{\theta}} \rightarrow \boldsymbol{\theta}$, $E(\mathbf{v}) \rightarrow \mathbf{0}$

四、最大似然估计方法——Newton-Raphson算法



Recall:

$$\Delta \hat{\theta} = - \left[\frac{\partial^2 J}{\partial \theta \partial \theta^T} \Big|_{\theta=\theta_0} \right]^{-1} \frac{\partial J}{\partial \theta} \Big|_{\theta=\theta_0}$$

参数迭代增量 $\Delta \hat{\theta}$

$$\Delta \hat{\theta} = \left[\sum_{i=1}^N \frac{\partial y^T(i)}{\partial \theta} \hat{R}^{-1} \frac{\partial y(i)}{\partial \theta} \right]^{-1} \Big|_{\theta=\theta_0} \left[\sum_{i=1}^N \frac{\partial y^T(i)}{\partial \theta} \hat{R}^{-1} v(i) \right] \Big|_{\theta=\theta_0}$$

定义 $\frac{\partial y}{\partial \theta}$ 为输出灵敏度矩阵，表征参数 θ 变化对输出 y 的影响

定义 Fisher信息矩阵 M

$$M = E \left[\frac{\partial J^2(\theta)}{\partial \theta \partial \theta^T} \right] = \sum_{i=1}^N \frac{\partial y^T(i)}{\partial \theta} \hat{R}^{-1} \frac{\partial y(i)}{\partial \theta}$$

最大似然估计参数

$$\hat{\theta} = \theta_0 - M_{\theta=\theta_0}^{-1} \frac{\partial J}{\partial \theta} \Big|_{\theta=\theta_0} \quad \text{Cov}(\hat{\theta}) \geq M_{\theta=\hat{\theta}}^{-1}$$

估计参数协方差的下界

四、最大似然估计方法——Newton-Raphson算法



一般采用数值计算方法计算输出灵敏度矩阵 $\frac{\partial y}{\partial \theta}$

$$\frac{\partial y}{\partial \theta_j} = \frac{y(\theta_0 + \delta\theta_j) - y(\theta_0)}{|\delta\theta_j|}$$

$$j = 1, 2, \dots, n_p$$

$$\frac{\partial y}{\partial \theta_j} = \frac{y(\theta_0 + \delta\theta_j) - y(\theta_0 - \delta\theta_j)}{2|\delta\theta_j|}$$



四、最大似然估计方法——Newton-Raphson算法

注意：

- 假设 θ_0 逼近于的 θ 真值， θ_0 可以由最小二乘方法获得；
- 每计算一步 $\Delta\hat{\theta}$ ，更新一次 \hat{R}

截止条件：

1. $\Delta\hat{\theta}$ 的绝对值足够小；
2. 两次迭代更新之间的代价函数 $J(\hat{\theta})$ 的变化足够小；
3. 代价函数梯度 $\frac{\partial J(\theta)}{\partial \theta}$ 足够接近于0；
4. \hat{R} 的变化足够小。

四、最大似然估计方法——Newton-Raphson算法



截止条件：

$\Delta\hat{\theta}$ 的绝对值足够小；

$$\left| (\hat{\theta}_j)_k - (\hat{\theta}_j)_{k-1} \right| < 1.0 \times 10^{-5}$$

或

$$\frac{\|\hat{\theta}_k - \hat{\theta}_{k-1}\|}{\hat{\theta}_{k-1}} < 0.001$$

同时， 两次迭代更新之间的代价函数 $J(\hat{\theta})$ 的变化足够小；

$$\left| \frac{J(\hat{\theta}_k) - J(\hat{\theta}_{k-1})}{J(\hat{\theta}_{k-1})} \right| < 0.001$$

四、最大似然估计方法——Newton-Raphson算法



截止条件：

同时，代价函数梯度 $\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ 足够接近于**0**；

$$\left| \left(\frac{\partial J(\boldsymbol{\theta})}{\partial \theta_j} \right)_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_k} \right| < 0.05$$

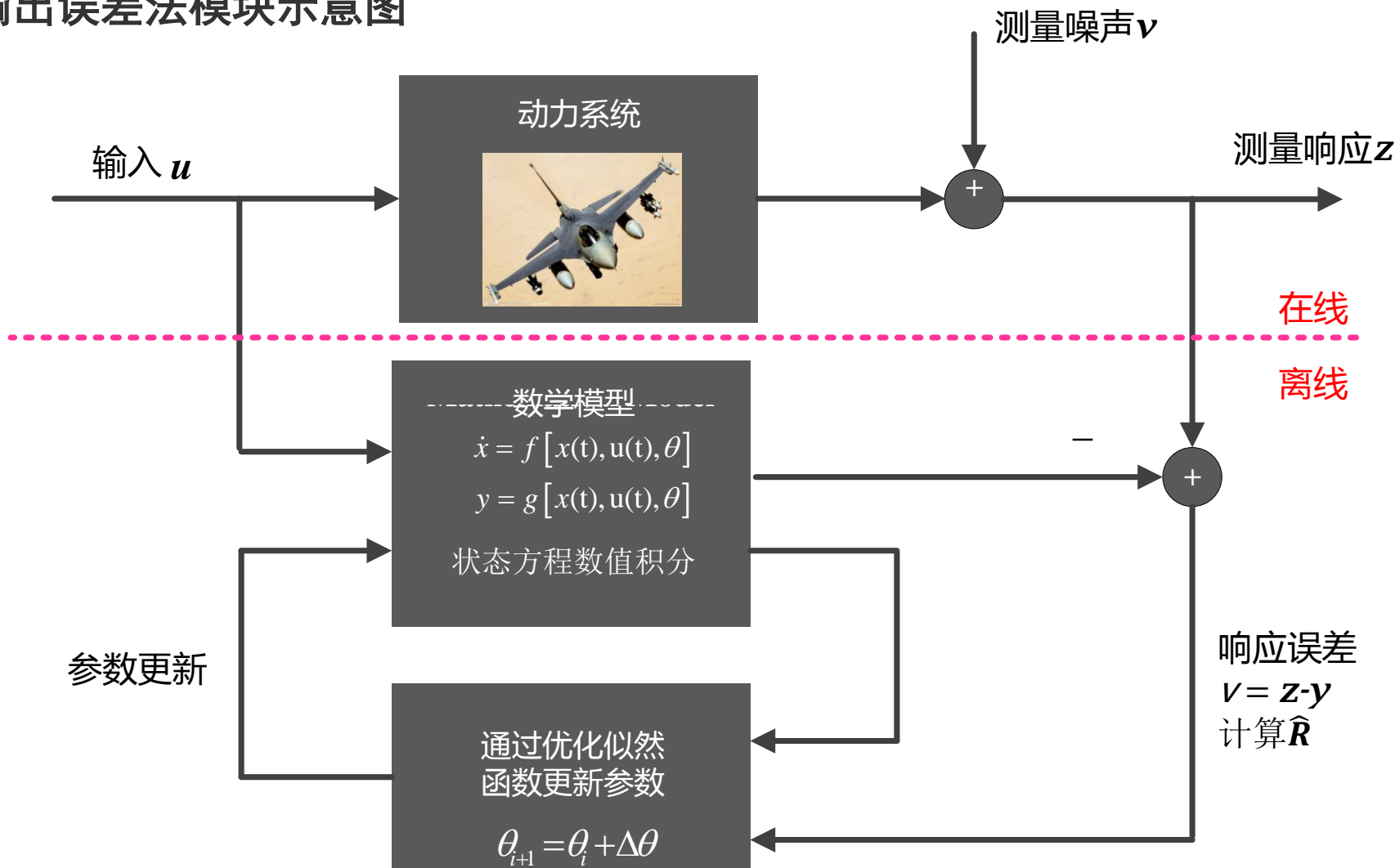
同时， $\hat{\mathbf{R}}$ 的变化足够小

$$\left| \frac{(\hat{r}_{jj})_k - (\hat{r}_{jj})_{k-1}}{(\hat{r}_{jj})_{k-1}} \right| < 0.05$$

四、最大似然估计方法——Newton-Raphson算法



输出误差法模块示意图



四、最大似然估计方法——例子

实例：NASA双水獭飞机侧向机动

动力学模型方程：

$$\dot{\beta} = \frac{\bar{q}_0 S}{m V_0} \left(C_{Y_\beta} \beta + C_{Y_r} \frac{rb}{2V_0} + C_{Y_{\delta_r}} \delta_r \right) + p \sin \alpha_0 - r \cos \alpha_0 + \frac{g \cos \theta_0}{V_0} \phi + b_{\dot{\beta}}$$

$$I_x \dot{p} - I_{xz} \dot{r} = \bar{q}_0 S b \left(C_{l_\beta} \beta + C_{l_p} \frac{pb}{2V_0} + C_{l_r} \frac{rb}{2V_0} + C_{l_{\delta_a}} \delta_a + C_{l_{\delta_r}} \delta_r \right) + b_{\dot{p}}$$

$$I_z \dot{r} - I_{xz} \dot{p} = \bar{q}_0 S b \left(C_{n_\beta} \beta + C_{n_p} \frac{pb}{2V_0} + C_{n_r} \frac{rb}{2V_0} + C_{n_{\delta_a}} \delta_a + C_{n_{\delta_r}} \delta_r \right) + b_{\dot{r}}$$

$$\dot{\phi} = p + \tan \theta_0 r + b_{\dot{\phi}}$$

$$a_y = \frac{\bar{q}_0 S}{mg} \left(C_{Y_\beta} \beta + C_{Y_r} \frac{rb}{2V_0} + C_{Y_{\delta_r}} \delta_r \right) + b_{a_y}$$



四、最大似然估计方法——例子

实例：NASA双水獭飞机侧向机动

系统方程：

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -I_{xx} & -I_{xz} & 0 \\ 0 & -I_{xz} & -I_{zz} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{\beta} \\ \dot{p} \\ \dot{r} \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} k_1 C_{Y_\beta} & \sin \alpha_0 & (k_1 C_{Y_r} k_2 - \cos \alpha_0) & \frac{g \cos \theta_0}{V_0} \\ k_3 C_{l_\beta} & k_3 C_{l_p} k_2 & k_3 C_{l_r} k_2 & 0 \\ k_3 C_{n_\beta} & k_3 C_{n_p} k_2 & k_3 C_{n_r} k_2 & 0 \\ 0 & 1 & \tan \theta_0 & 0 \end{bmatrix} \begin{bmatrix} \beta \\ p \\ r \\ \phi \end{bmatrix} + \begin{bmatrix} 0 & k_1 C_{Y_{\delta_r}} & b_{\dot{\beta}} \\ k_3 C_{l_{\delta_a}} & k_3 C_{l_{\delta_r}} & b_{\dot{p}} \\ k_3 C_{n_{\delta_a}} & k_3 C_{n_{\delta_r}} & b_{\dot{r}} \\ 0 & 0 & b_{\dot{\phi}} \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_r \\ 1 \end{bmatrix}$$

测量方程：

$$\begin{bmatrix} \beta \\ p \\ r \\ \phi \\ a_y \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ k_4 C_{Y_\beta} & 0 & k_4 C_{Y_r} k_2 & 0 \end{bmatrix} \begin{bmatrix} \beta \\ p \\ r \\ \phi \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & k_4 C_{Y_{\delta_r}} & b_{a_y} \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_r \\ 1 \end{bmatrix}$$

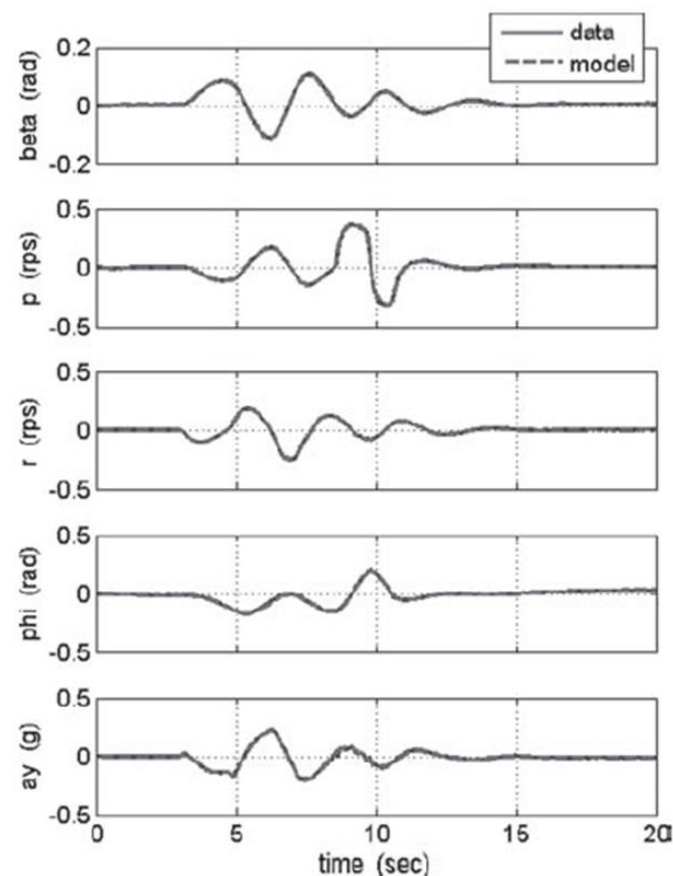
估计参数：

$$\theta = \left[C_{Y_\beta} \quad C_{Y_r} \quad C_{Y_{\delta_r}} \quad b_{\dot{\beta}} \quad C_{l_\beta} \quad C_{l_p} \quad C_{l_r} \quad C_{l_{\delta_a}} \quad C_{l_{\delta_r}} \quad b_{\dot{\beta}}' \quad C_{n_\beta} \quad C_{n_p} \quad C_{n_r} \quad C_{n_{\delta_a}} \quad C_{n_{\delta_r}} \quad b_{\dot{r}}' \quad b_{\dot{\phi}} \quad b_{a_y} \right]^T$$

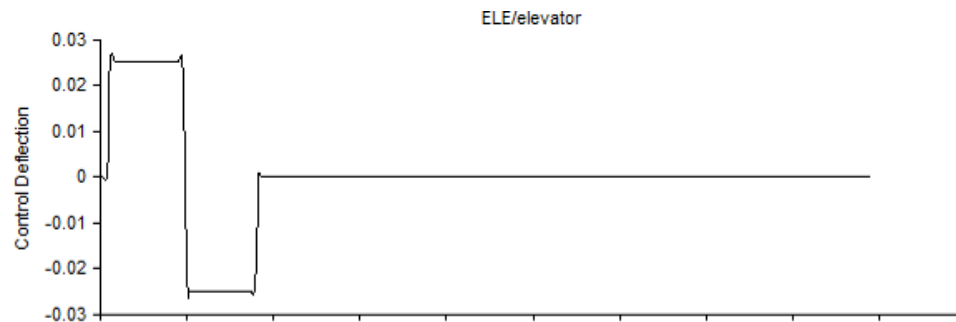
四、最大似然估计方法——例子

实例：NASA双水獭飞机侧向机动

Parameter	Output error		Equation error	
	$\hat{\theta}$	$s(\hat{\theta})$	$\hat{\theta}$	$s(\hat{\theta})$
$C_{Y\beta}$	-8.66×10^{-1}	2.80×10^{-3}	-8.45×10^{-1}	7.41×10^{-3}
C_{Yr}	9.31×10^{-1}	1.65×10^{-2}	9.63×10^{-1}	5.53×10^{-2}
$C_{Y\delta_r}$	3.75×10^{-1}	7.98×10^{-3}	3.43×10^{-1}	2.89×10^{-2}
$C_{l\beta}$	-1.19×10^{-1}	5.60×10^{-4}	-1.05×10^{-1}	2.54×10^{-3}
C_{lp}	-5.84×10^{-1}	2.39×10^{-3}	-5.28×10^{-1}	9.93×10^{-3}
C_{lr}	1.88×10^{-1}	1.71×10^{-3}	1.93×10^{-1}	8.33×10^{-3}
$C_{l\delta_a}$	-2.28×10^{-1}	8.43×10^{-4}	-2.12×10^{-1}	3.39×10^{-3}
$C_{l\delta_r}$	3.84×10^{-2}	7.12×10^{-4}	2.99×10^{-2}	3.86×10^{-3}
$C_{n\beta}$	8.65×10^{-2}	2.07×10^{-4}	8.54×10^{-2}	3.58×10^{-4}
C_{np}	-6.39×10^{-2}	9.49×10^{-4}	-5.15×10^{-2}	1.43×10^{-3}
C_{nr}	-1.92×10^{-1}	6.07×10^{-4}	-1.98×10^{-1}	1.30×10^{-3}
$C_{n\delta_a}$	-2.73×10^{-3}	3.50×10^{-4}	2.34×10^{-3}	5.00×10^{-4}
$C_{n\delta_r}$	-1.36×10^{-1}	2.66×10^{-4}	-1.31×10^{-1}	5.97×10^{-4}



五、时域辨识试验设计



Ail/aileron



获得包含某种动态特性的实验数据

五、时域辨识试验设计——输入信号设计



系统辨识实验的中心思想：**独立地、足够地**激发飞机运动相关的模态

定义一个输入的4个维度：激励时长、涉及的控制面、输入的波形、输入的幅值。

评价激励信号的标准：**非参数量度**：信噪比、相干性（coherence，频域）
参数化量度：输出灵敏度

好的建模结果对信号的要求：**信噪比大于10，相干性大于0.8**

输入的实际限制：1.安全

2.幅值的限制：
攻角与侧滑角：**正负 5 deg**
机体角速率：**正负 20 deg/s**
机体加速度：**正负 0.1g – 0.3g**



幅值设计:

输入信号幅值不宜过大或过小。

幅值过大, 会使系统进入非线性区域; 而幅值过小, 又会使系统中的噪声占主导地位

采样频率的选择:

1. 满足采样定理, 采样频率不低于信号截至频率的两倍, 即采样间隔 T_0 不宜太大, 否则信息量损失大
2. T_0 也不宜太小, T_0 太小时, 由于计算机的截断误差, $k+1$ 时刻的数据和 k 时刻的数据可能没有差别。

数据长度的选择:

考虑多方面因素, 且采集的数据长度与感兴趣的模态周期相匹配

五、时域辨识试验设计——输入信号设计

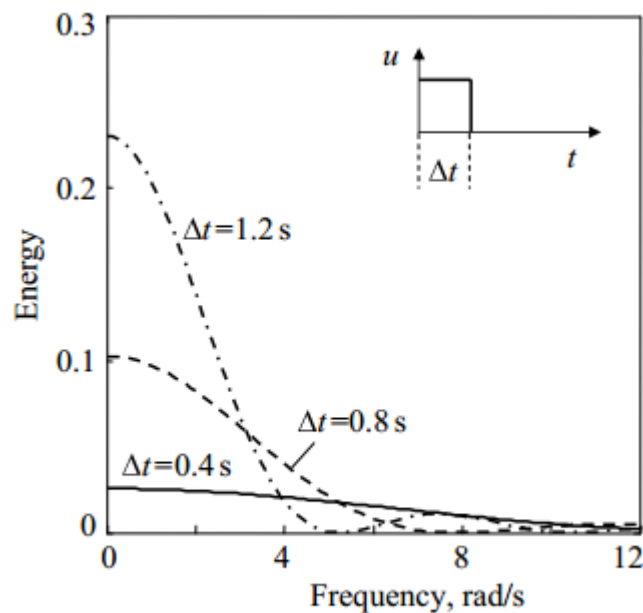
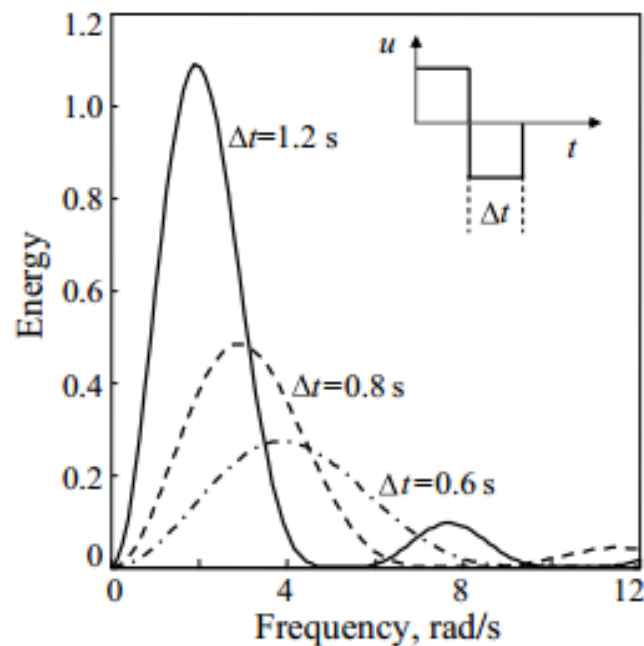


Fig. 2.5 Energy spectra of pulse inputs.

脉冲输入：不适用于激励稍高频率的快速模态

如：纵向短周期与荷兰滚模态。



Doublet：2-sided pulses
正弦波的方波近似

Doublet 的脉冲时间选择要与动态系统的主导频率相近

$$\Delta t_{\text{DBLT}} \approx \frac{2.3}{\omega_n} \approx \frac{2\pi}{2.7\omega_n} \approx \frac{1}{2.7} \cdot \text{period of oscillation}$$

五、时域辨识试验设计——输入信号设计



Multistep : 3-2-1-1

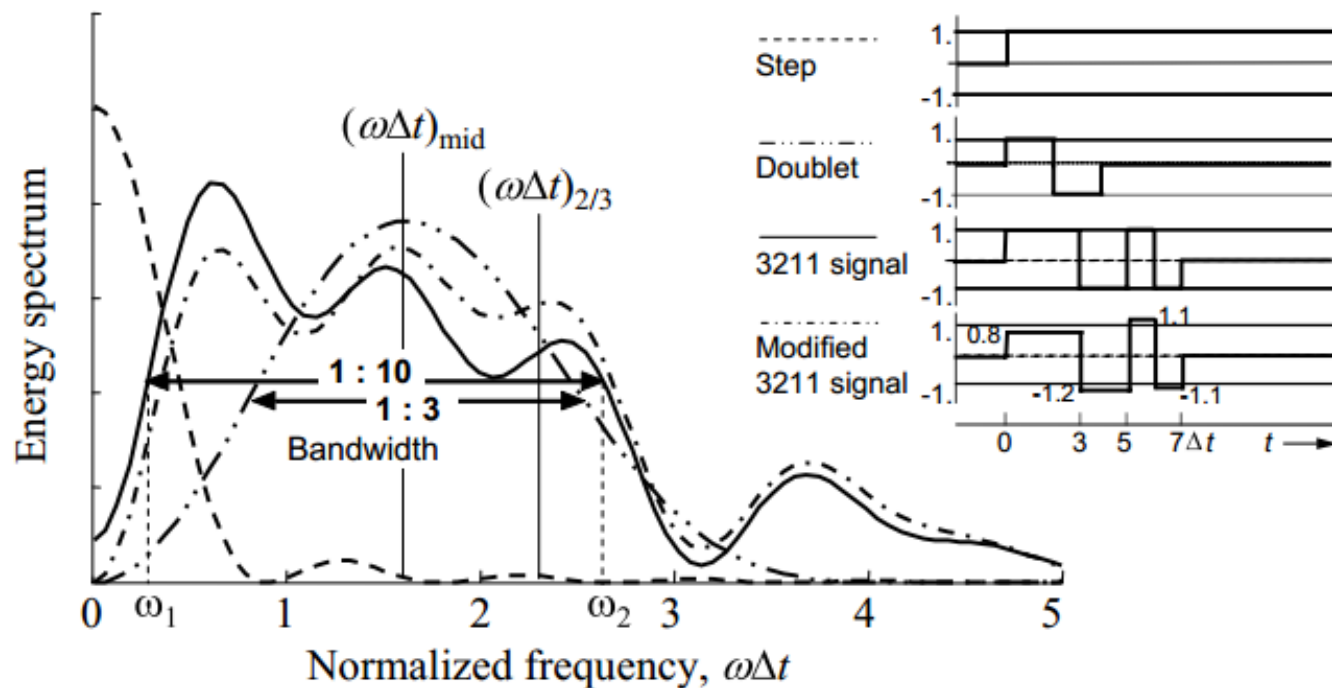


Fig. 2.7 Frequency domain comparison of standard inputs.

$$\Delta t_{3211} \approx \frac{2.1}{\omega_n} \approx \frac{1}{3} \cdot \text{period of oscillation}$$

最典型、最常用的时域信号输入：“穷人的扫频”

典型的激励信号：

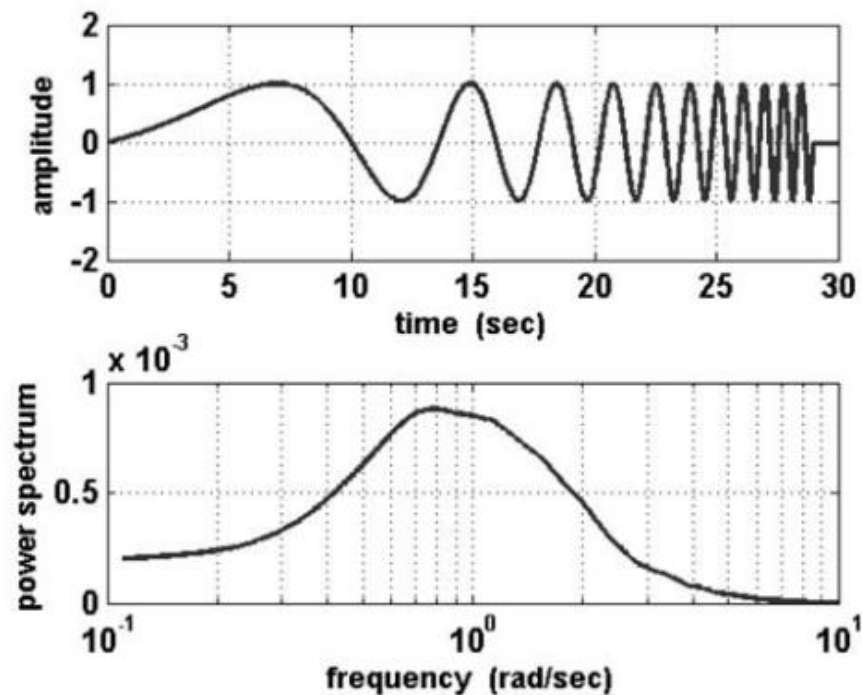


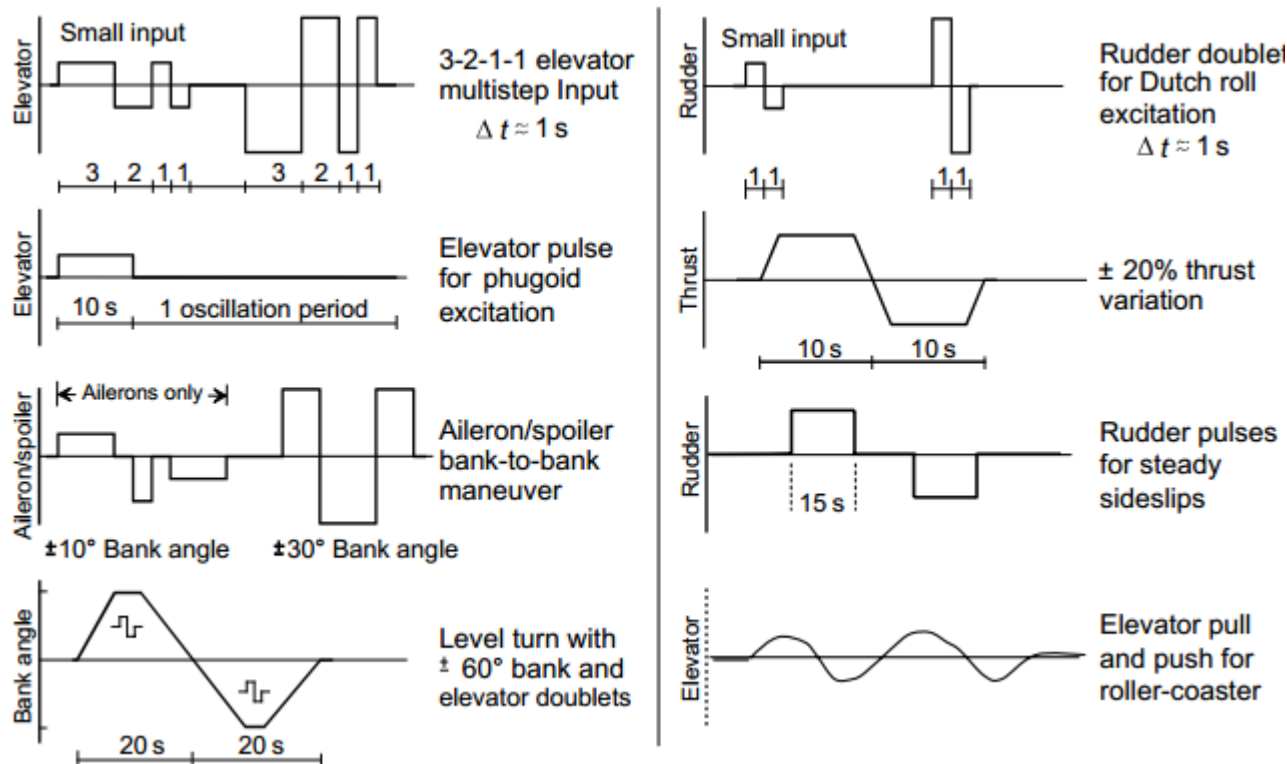
扫频信号： 随时间的增加频率逐渐加快的连续正弦信号

要求激励的起始与终止信号的幅值都为零，扫频信号先激励的是系统的低频内容，后是高频段的内容。

应用时无需系统的先验信息，综合覆盖频率段扫频得到模型一般具有良好的预测能力；

用时长（60秒-90秒），
对于多输入问题效率低下，
扫频信号有可能导致结构共振



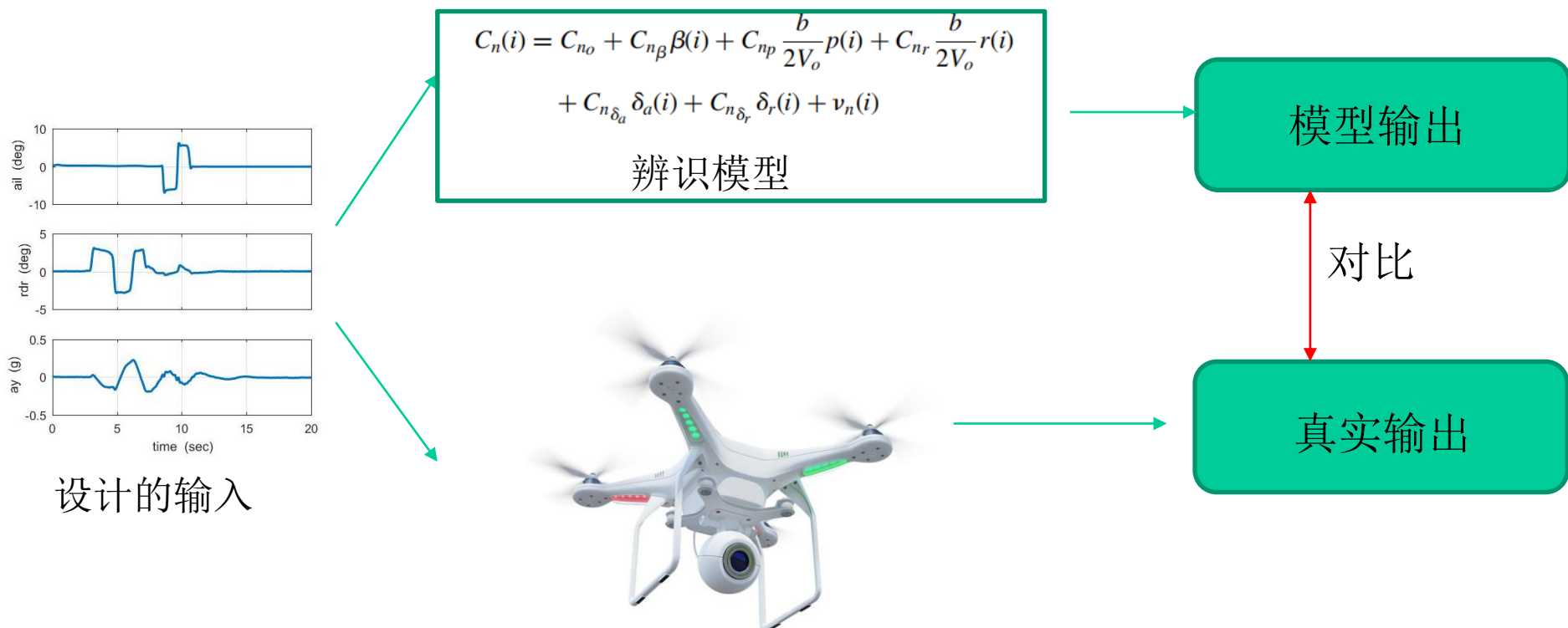


典型输入应用场景：纵向短周期模态; 横侧向荷兰滚模态、快速滚转模态

五、时域辨识试验设计——模型验证



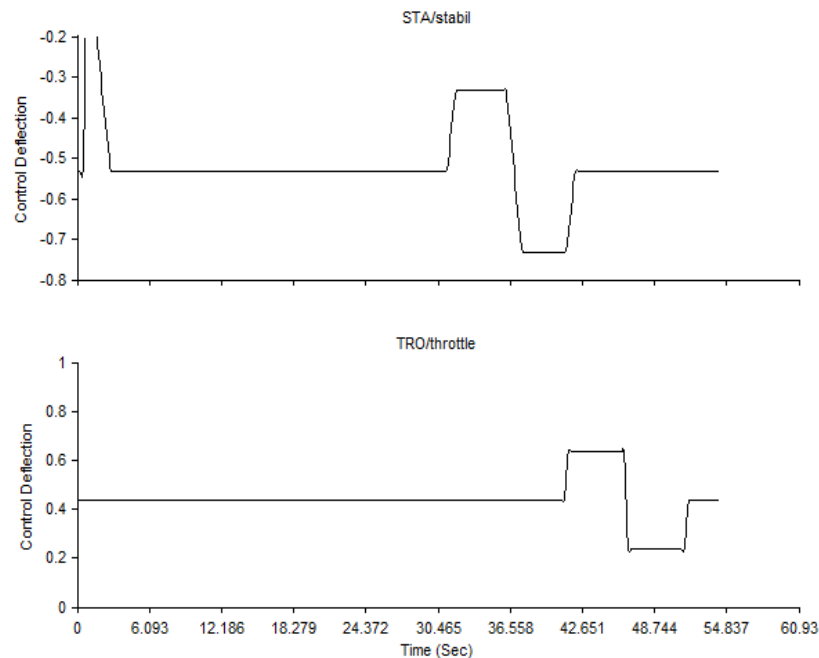
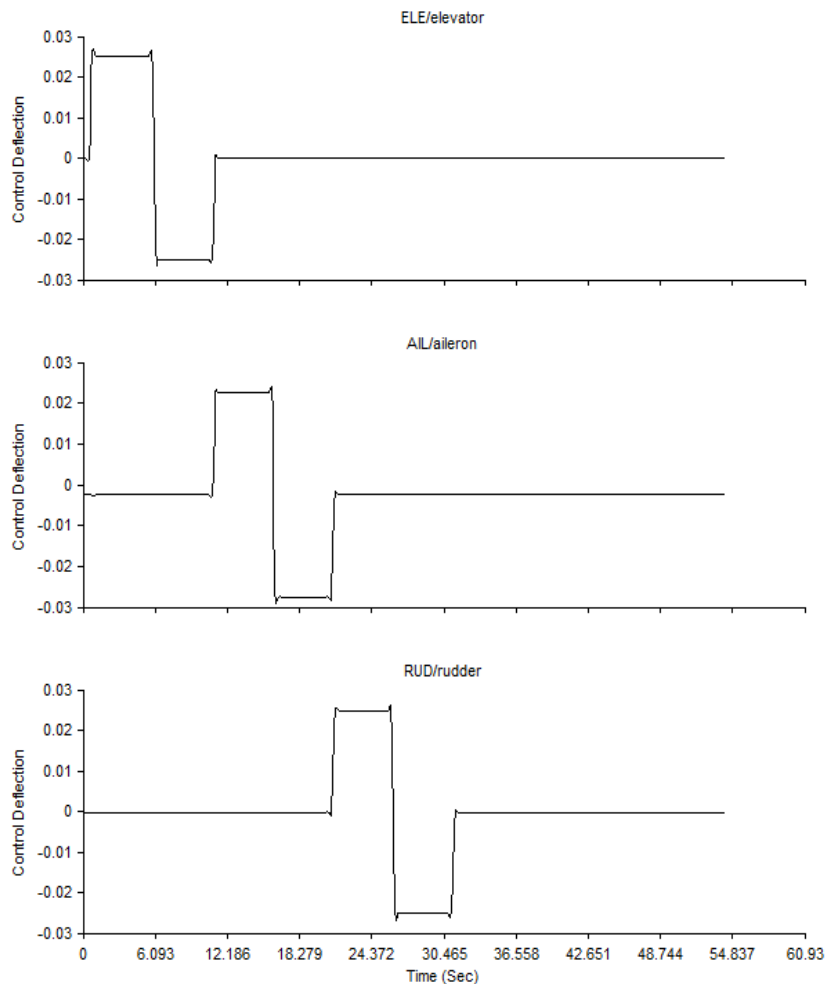
对比模型输出和真实输出，分析模型的预测能力



五、时域辨识试验设计——模型验证



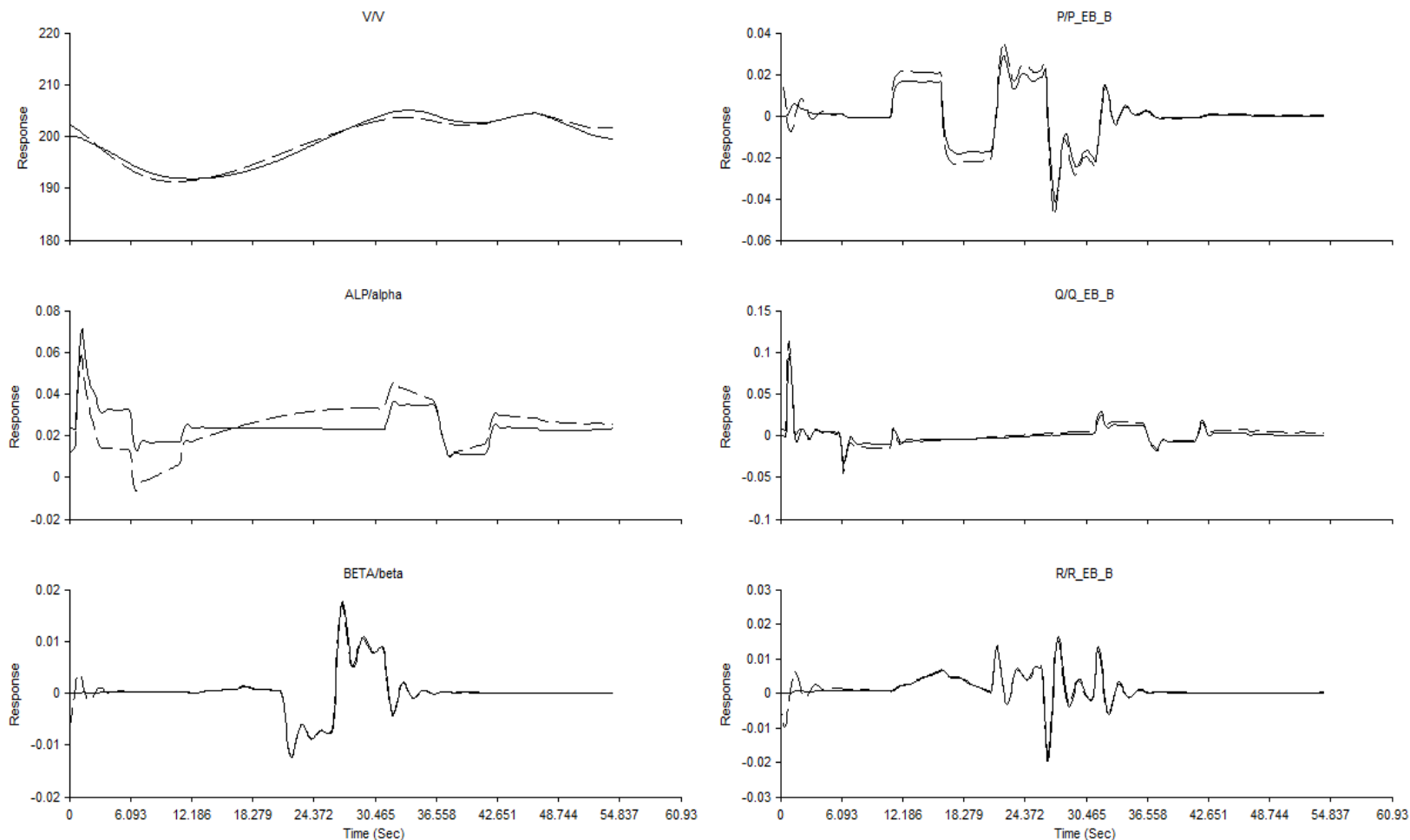
输入:



五、时域辨识试验设计——模型验证



模型输出和真实输出:



The background of the slide is a photograph of a flight test. Several aircraft are visible in the sky, each leaving a long, curved trail of white smoke. The trails are arranged in a pattern that suggests a coordinated maneuver. The sky is filled with clouds, and the sun is visible on the left side, creating a bright glow.

谢谢