Assignment 5: Model test and outlier removal

Kiran Sripathy, Simone Chellini

July 5, 2024

```
[ ]: import numpy as np
     from numpy.linalg import inv
     from matplotlib import pyplot as plt
```

$y = y_0 + a\sin(2\pi f_0 t)$

$f_0 = 1$

$y = y_0 + a\sin(2\pi t)$

$t_= \begin{bmatrix} 0.25 & 0.50 & 0.75 & 1.00 \end{bmatrix}^T$

$y = \begin{bmatrix} 5.12 & 5.04 & 4.90 & 5.02 \end{bmatrix}^T$

0.1 5-a

$$A = \begin{bmatrix} 1 & \sin(0.5\pi) \\ 1 & \sin(\pi) \\ 1 & \sin(1.5\pi) \\ 1 & \sin(2\pi) \end{bmatrix}$$

$m = 4$

$n = 2$

Redundancy is $m - n = 2$

0.2 5-b

```
[ ]: y = np.array([[5.12],[5.04],[4.90],[5.02]])
     t = np.array([[0.25],[0.50],[0.75],[1.00]])

     A0 = np.array([1,1,1,1])
     A1 = np.sin(2*np.pi*t)
     A = np.column_stack((A0,A1))

     x_hat_ls = inv(A.T @ A) @ A.T @ y
```

$\hat{x} = \begin{bmatrix} 5.02 & 0.11 \end{bmatrix}^T$

$\hat{y}_0 = 5.02; \hat{a} = 0.11$

0.3 5-c

$$Q_y = \begin{bmatrix} 0.01^2 & 0 & 0 & 0 \\ 0 & 0.01^2 & 0 & 0 \\ 0 & 0 & 0.01^2 & 0 \\ 0 & 0 & 0 & 0.01^2 \end{bmatrix}$$

[ ]:
```
Qy = np.array([[0.01**2,0,0,0],[0,0.01**2,0,0],[0,0,0.01**2,0],[0,0,0,0.01**2]])

x_hat_blue = inv(A.T @ inv(Qy) @ A) @ A.T @ inv(Qy) @ y
```

$\hat{x} = \begin{bmatrix} 5.02 & 0.11 \end{bmatrix}^T$

$\hat{y}_0 = 5.02; \hat{a} = 0.11$

0.4 5-d

$Q_{\hat{x}} = (A^T Q_y^{-1} A)^{-1}$

[ ]:
```
Qx = inv(A.T @ inv(Qy) @ A)
sigma_x = np.sqrt(np.diag(Qx))
```

$$Q_{\hat{x}} = \begin{bmatrix} 2.5 \times 10^{-5} & 0 \\ 0 & 5 \times 10^{-5} \end{bmatrix}^T$$

$\sigma_{\hat{y}_0} = 0.005$

$\sigma_{\hat{a}} = 0.0071$

0.5 5-e

$y(t_5 = 1.25) = y_5 = 5.02 + 0.11 \times \sin(2.5\pi)$

[ ]:
```
t5 = 1.25
y5_hat = x_hat_blue[0] + x_hat_blue[1]*np.sin(2*np.pi*t5)
```

$\hat{y}_5 = 5.13$

$A_{5j} = \begin{bmatrix} 1 & \sin 2.5\pi \end{bmatrix}$

Since all the emasurements are uncorrelated:

$Q_{\hat{y}_{55}} = A_{5j} Q_{\hat{x}} A_{5j}^T$

[ ]:
```
A5j = np.array([1,np.sin(2*np.pi*t5)])
Qy5_hat = A5j @ Qx @ A5j.T
sigma_y5 = np.sqrt(Qy5_hat)
```

$Q_{\hat{y}_{55}} = 7.5 \times 10^{-5}$

$\sigma_{\hat{y}_5} = 0.00866$

0.6 5-f

$y(t=1.50)=5.02+0.11\times\sin(3\pi)$

[ ]:
```
t6 = 1.50
y6_hat = x_hat_blue[0] + x_hat_blue[1]*np.sin(2*np.pi*t6)
```

$y(t=1.50)=5.02$

$A_{6j}=\begin{bmatrix}1 & \sin 3.0\pi\end{bmatrix}$

Similarly:

$Q_{\hat{y}_{66}}=A_{6j}Q_{\hat{x}}A_{6j}^T$

[ ]:
```
A6j = np.array([1,np.sin(2*np.pi*t6)])
Qy6_hat = A6j @ Qx @ A6j.T
sigma_y6 = np.sqrt(Qy6_hat)
```

$\sigma_{\hat{y}_6}=0.005$

0.7 5-g

$\hat{y}=A\hat{x}$

[ ]:
```
y_hat = A @ x_hat_blue
```

$\hat{y}=\begin{bmatrix}5.13 & 5.02 & 4.91 & 5.02\end{bmatrix}^T$

$\hat{e}=y-\hat{y}$

[ ]:
```
e_hat = y - y_hat
```

$\hat{e}=\begin{bmatrix}-0.01 & 0.02 & -0.01 & 0\end{bmatrix}^T$

$\|\hat{e}\|^2_{Q_y^{-1}}=\hat{e}^TQ_y^{-1}\hat{e}$

[ ]:
```
weighted_squared_e_hat = e_hat.T @ inv(Qy) @ e_hat
```

$\|\hat{e}\|^2_{Q_y^{-1}}=6$

0.8 5-h

$\|\hat{e}\|^2_{Q_y^{-1}}\sim\chi^2(m-n)=\chi^2(2)$

0.9 5-i

For $\chi^2(2); k_{0.01}=9.2103$.

0.10 5-j

$\|\hat{e}\|^2_{Q_y^{-1}} < k_{0.01}$

The alternate hypothesis that the assumed model $y = y_0 + a\sin(2\pi f_0 t)$ not being the right fit to the measured data can be rejected with 99% confidence level.

0.11 5-k

$Q_{\hat{e}} = Q_y - Q_{\hat{y}}$

[ ]:
```
Qy_hat = A @ Qx @ A.T
```

$$Q_{\hat{y}} = \begin{bmatrix} 7.5\times10^{-5} & 2.5\times10^{-5} & -2.5\times10^{-5} & 2.5\times10^{-5} \\ 2.5\times10^{-5} & 2.5\times10^{-5} & 2.5\times10^{-5} & 2.5\times10^{-5} \\ -2.5\times10^{-5} & 2.5\times10^{-5} & 7.5\times10^{-5} & 2.5\times10^{-5} \\ 2.5\times10^{-5} & 2.5\times10^{-5} & 2.5\times10^{-5} & 2.5\times10^{-5} \end{bmatrix}$$

[ ]:
```
Q_e_hat = Qy - Qy_hat
sigma_e_hat = np.sqrt(np.diag(Q_e_hat))
```

$$Q_{\hat{e}} = \begin{bmatrix} 2.5\times10^{-5} & -2.5\times10^{-5} & 2.5\times10^{-5} & -2.5\times10^{-5} \\ -2.5\times10^{-5} & 7.5\times10^{-5} & -2.5\times10^{-5} & -2.5\times10^{-5} \\ 2.5\times10^{-5} & -2.5\times10^{-5} & 2.5\times10^{-5} & -2.5\times10^{-5} \\ -2.5\times10^{-5} & -2.5\times10^{-5} & -2.5\times10^{-5} & 7.5\times10^{-5} \end{bmatrix}$$

$\sigma_{\hat{e}_1} = 0.005; \sigma_{\hat{e}_2} = 0.00866; \sigma_{\hat{e}_3} = 0.005; \sigma_{\hat{e}_4} = 0.00866$

0.12 5-l

$w_i = \frac{\hat{e}_i}{\sigma_{\hat{e}_i}}$

[ ]:
```
norm_e_hat = e_hat.reshape(-1)/sigma_e_hat.reshape(-1)
```

$w_1 = -2; w_2 = 2.309; w_3 = -2; w_4 = 0$

Since the standard deviation of the measurements are known, $\underline{w}_i \sim N(0,1)$.

0.13 5-m

For 1% significance level and $N(0,1), k_{0.005} = 2.575$.

0.14 5-n

The critical region for outlier detection are $k_\alpha < -2.575$ and $k_\alpha > 2.575$. Normalized residuals for all the four observations fall outside the critical region. Therefore none of the observations are outliers at 99% confidence level.