ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

Author - Aloy Banerjee

Roll No. CH22M503

Part 1

Part 1

Q1

In [4]:
```
tk = np.array([0, 1, 2, 3, 4])
yk = np.array([1.87, 5.4, 6.86, 12.76, 19.89])
```

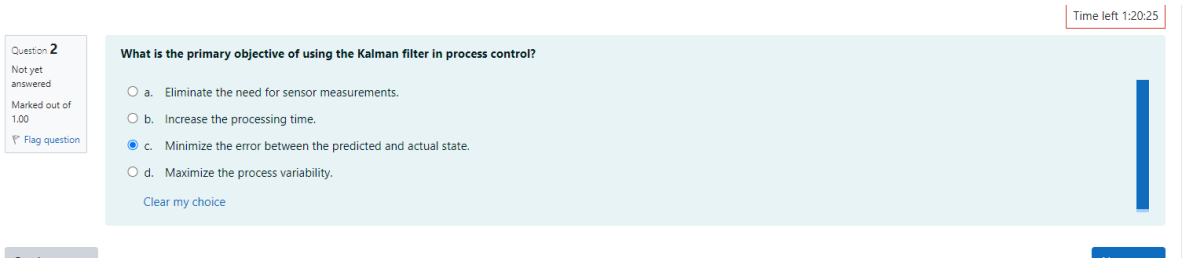
In [14]:
```
def design_matrix(x, L=1):
    return np.column_stack((np.ones(len(x)), x))

def predicted_values(x, theta):
    X = design_matrix(x)
    return X @ theta

def true_values(x, theta):
    X = design_matrix(x)
    return X @ theta
```

In [15]:
```
x_observed = tk
y_observed = yk
```

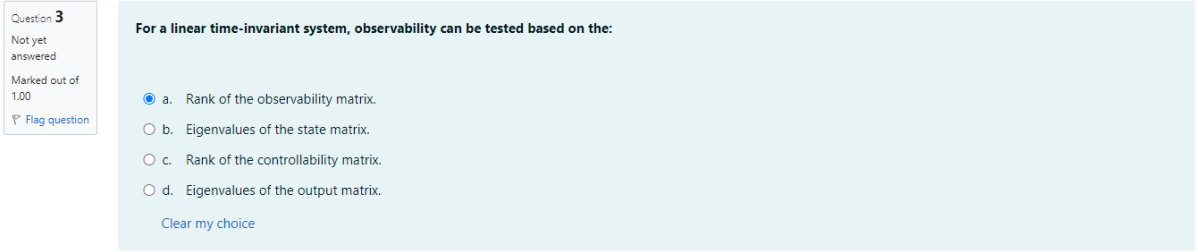
In [16]:
```
theta_estimated = np.linalg.lstsq(design_matrix(x_observed,L), y_observed, rcond=None)[0]
```

In [17]:
```
theta_estimated
```

Out[17]:
```
array([0.676, 4.34 ])
```



The Kalman filter is used to estimate a system's state based on noisy observations, with the goal of minimising the difference between the predicted and actual states. This makes it particularly effective in process control applications where the system state cannot be directly measured or is subject to severe noise. It is not used to reduce sensor measurements, speed up processing, or maximise process variability.



ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

Author - Aloy Banerjee

Roll No. CH22M503

Observability is a measure of how well internal states of a system can be inferred by knowledge of its external outputs. For a linear time-invariant (LTI) system, the system is observable if the observability matrix is full rank. This is because a full-rank observability matrix ensures that the current state of the system can be determined using output data from the system over time.

Note: Eigenvalues of the state matrix are used to determine the stability of the system, not its observability. The rank of the controllability matrix is used to check the controllability (i.e., whether the system's state can be driven from any initial state to any final state in finite time), not observability. Eigenvalues of the output matrix don't directly provide information about observability.

Question 4
Not yet answered
Marked out of 1.00
Flag question

For a system to be controllable and observable, the controllability and observability matrices must be:

- a. Identity matrices.
- b. Symmetric matrices.
- c. Full rank square matrices.
- d. Diagonal matrices

Clear my choice

For a system to be controllable, the controllability matrix must be full rank. Similarly, for a system to be observable, the observability matrix must be full rank. It's not required for these matrices to be identity, symmetric, or diagonal matrices. The full rank condition ensures that each state of the system can be independently influenced (controllability) and observed (observability).

Question 5
Not yet answered
Marked out of 1.00
Flag question

The Kalman filter assumes that the process and measurement noises are:

- a. Correlated and have non-zero mean.
- b. Uncorrelated and have non-zero mean.
- c. Uncorrelated and have zero mean.
- d. Correlated and have zero mean.

Clear my choice

The Kalman filter assumes that the process and measurement noises are uncorrelated (i.e., independent of each other) and have a zero mean (i.e., they are white noise). This assumption is fundamental to its operation. It also typically assumes that the noises are Gaussian (normally

ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

Author - Aloy Banerjee

Roll No. CH22M503

distributed), although the filter can still provide reasonable estimates under non-Gaussian conditions.

Question 6
Not yet answered
Marked out of 1.00
Flag question

In the Kalman filter, the prediction step involves:

- a. Estimating the aposteriori process covariance matrix.
- b. Computing the Kalman gain.
- c. Combining the prediction and measurement updates.
- d. Projecting the state ahead of time. (selected)

Clear my choice

In the Kalman filter, the prediction step involves using the system dynamics to project the current state and covariance estimates ahead in time. This means predicting the state of the system at the next time step based on the system model and the current state.

The a posteriori process covariance matrix estimation and computing the Kalman gain are part of the update (or correction) step, not the prediction step. Similarly, combining the prediction and measurement updates is also part of the update step.

Question 7
Not yet answered
Marked out of 1.00
Flag question

The Kalman filter is most efficient when the process and measurement noise distributions are:

- a. Poisson.
- b. Uniform.
- c. Gaussian. (selected)
- d. Exponential.

Clear my choice

The Kalman filter is optimal and most efficient when the process and measurement noise distributions are Gaussian (or Normal). This is because the Kalman filter algorithm involves procedures (like weighted averaging) that are optimal under the least mean square error criterion when the system is linear, and the noise is Gaussian.

ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

Author - Aloy Banerjee

Roll No. CH22M503

While the filter can still provide reasonable estimates with non-Gaussian noise, its performance in terms of accuracy and efficiency can be significantly reduced.

Question **8**

Not yet answered

Marked out of 1.00

Flag question

The process covariance matrix in the Kalman filter represents

- a. The uncertainty in the measurement model.
- b. The uncertainty in the sensor measurements.
- c. The uncertainty in the initial state estimate.
- d. The uncertainty in the process model. (selected)

Clear my choice

The process covariance matrix, often denoted as Q, in the Kalman filter represents the uncertainty or noise in the process model. It models the statistical properties of the process noise, that is, how much noise or uncertainty is introduced into the state of the system through the process itself.

Note that the uncertainty in the measurement model is typically represented by the measurement covariance matrix (often denoted as R), while the uncertainty in the initial state estimate is represented in the initial state covariance matrix (often denoted as P_0). The uncertainty in the sensor measurements is also typically represented by the measurement covariance matrix R.

Question **9**

Not yet answered

Marked out of 1.00

Flag question

Which of the following statements is true about the Kalman filter's recursive nature?

- a. It assumes the process is deterministic.
- b. It requires storage of the entire measurement history.
- c. It uses past estimates to update the current estimate. (selected)
- d. It can only process time-invariant systems.

Clear my choice

The Kalman filter is recursive in nature, which means it uses the estimated state from the previous time step and the current measurements to produce a new estimate. This approach allows it to process information as it becomes available over time, rather than needing to store and reprocess all

ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

Author - Aloy Banerjee

Roll No. CH22M503

past measurements. The other options are not accurate: The Kalman filter doesn't assume the process is deterministic (it deals well with stochastic processes), it doesn't require storage of the entire measurement history (due to its recursive nature), and it isn't limited to time-invariant systems (though the standard Kalman filter is often applied to linear time-invariant systems, there are extensions for nonlinear and time-varying systems).

Question 10
Not yet answered
Marked out of 1.00
Flag question

In the Kalman filter, the measurement update step involves:

- a. Predicting the state ahead in time.
- b. Ignoring the measurement data.
- c. Computing the Kalman gain.
- d. Estimating the process covariance matrix.

Clear my choice

In the Kalman filter, the measurement update step (also known as the correction or innovation step) involves computing the Kalman gain. The Kalman gain determines the extent to which the measurement should be incorporated into the new state estimate.

During this step, the filter also updates the state estimate using the new measurement and the Kalman gain and updates the state covariance estimate. The process of predicting the state ahead in time is part of the prediction step, not the measurement update step. The Kalman filter does not ignore the measurement data; rather, it uses this data to refine the state estimate. And while the measurement update step does involve updating the state covariance estimate, it typically doesn't directly involve estimating the process covariance matrix (which is generally considered a known part of the system model).

Question 11
Not yet answered
Marked out of 1.00
Flag question

The Kalman filter is optimal when the process and measurement noise is:

- a. Non-Gaussian.
- b. Correlated.
- c. Gaussian and white.
- d. Heavily biased.

Clear my choice

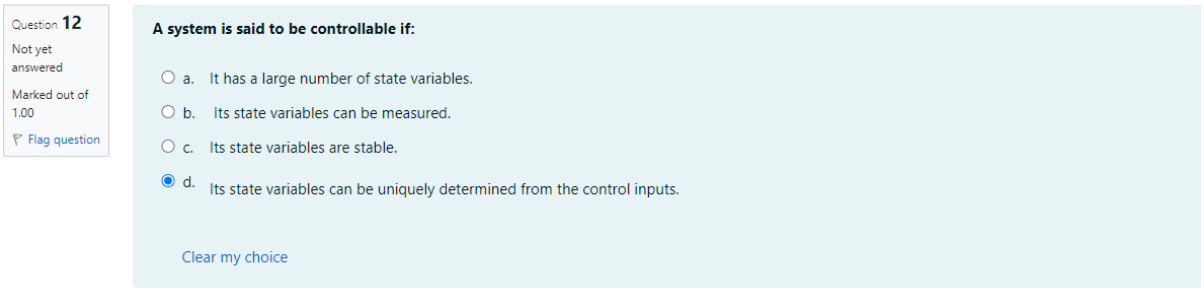
ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

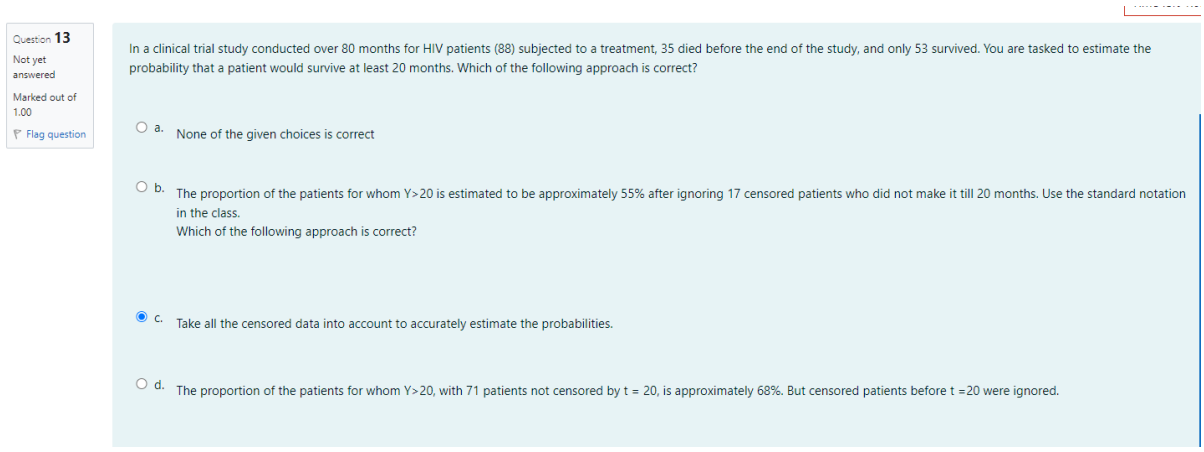
Author - Aloy Banerjee

Roll No. CH22M503

The Kalman filter is optimal when the process and measurement noise is Gaussian and white. "White" noise refers to noise that is uncorrelated over time, has zero mean, and has a constant power spectral density. It is also assumed that the process noise and measurement noise are uncorrelated with each other. These assumptions are important for the optimality of the Kalman filter in terms of minimizing the mean squared error of the state estimate. Non-Gaussian, correlated, or heavily biased noise could significantly degrade the performance of the filter.



A system is said to be controllable if any desired state of the system can be achieved by using appropriate control inputs within a finite time. In other words, the state variables can be driven to a particular value or condition based on the control inputs. This doesn't necessarily have anything to do with the number of state variables, the ability to measure those variables, or their stability.



Survival analysis is a branch of statistics for analysing the expected duration of time until one or more events happen, such as death in biological organisms. It involves the concept of "censoring" where the information about some individuals is incomplete.

ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

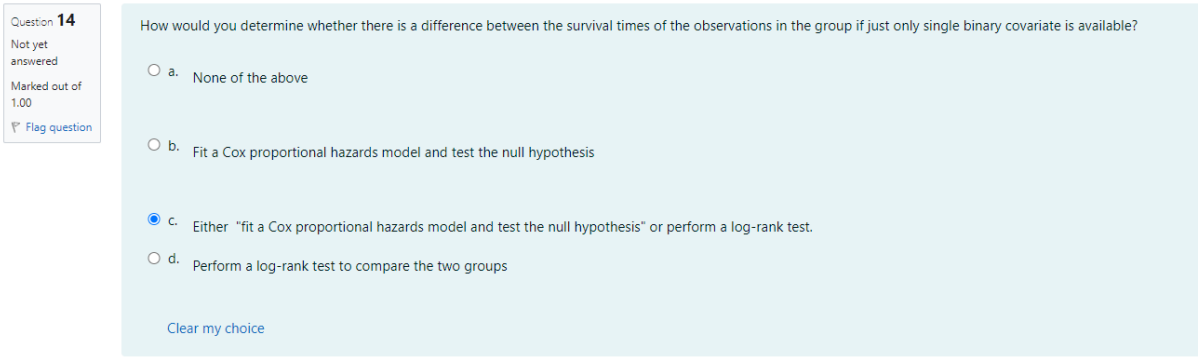
Author - Aloy Banerjee

Roll No. CH22M503

In a clinical trial study like the one described, if some patients did not die before the end of the study, their survival time would be censored because their exact survival time isn't known (only that it is longer than the duration of the study).

Ignoring these censored individuals in your analysis would bias your results, as it essentially assumes that these individuals would have died immediately after the end of the study. In contrast, taking all the censored data into account would give a more accurate estimation of the survival function and, therefore, the probability that a patient would survive at least 20 months.

To accomplish this, methods such as Kaplan-Meier estimator or the Cox proportional hazards model could be used, which are designed to handle censored data.



When comparing survival times between two groups categorized by a binary covariate, one has two methods at their disposal: the Cox proportional hazards model and the log-rank test.

The Cox proportional hazards model is a regression technique that allows you to study the impact of multiple variables on the time it takes for a specific event to occur. In survival analysis, this event is often related to mortality, making the Cox model particularly useful for studying survival scenarios. By fitting this model, one can test the null hypothesis, which examines whether the binary covariate (group indicator) has any significant effect on survival.

On the other hand, the log-rank test is a non-parametric hypothesis test utilized to compare the survival distributions of two or more groups. In this context, one would use it to investigate the null hypothesis, which questions whether there are any notable differences between the survival curves of the groups identified by the binary covariate.

Both methods can be suitable for analysis, and the decision on which one to use depends on the specific characteristics of the dataset aim to answer. If dataset comprises a single binary covariate and primary interest lies in determining if there's a difference between the defined groups, without focusing on the precise magnitude or nature of that difference, then the log-rank test could be a

ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

Author - Aloy Banerjee

Roll No. CH22M503

simpler option. On the contrary, if one seeks to control for potential confounding factors or are interested in quantifying the disparity between groups, the Cox model would be more appropriate in such cases.

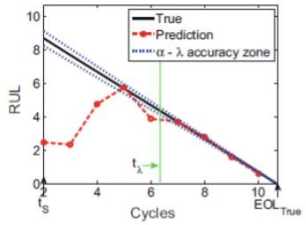
Question **15**

Not yet answered

Marked out of 2.00

Flag question

In an α-λ accuracy analysis for the predicted RUL from the 3rd cycle, λ is considered to be 0.5 and the true end of the life is 10.5 cycles. Which of the following statement is correct?



- a. α-λ accuracy is false
- b. α-λ accuracy cannot be determined.
- c. α-λ accuracy is true

Clear my choice

Question **16**

Answer saved

Marked out of 2.00

Flag question

Consider the covariance matrix for a process is given as follows

$$\begin{bmatrix} 1.01 & 0.25 & -0.02 \\ 0.025 & 1.01 & 0.04 \\ -0.017 & 0.04 & 1.01 \end{bmatrix}$$. The T^2 statistics for a new data point is $$\begin{bmatrix} 0.724 \\ 0.649 \\ 1.857 \end{bmatrix}$$

Answer: 1.6237874863769912

ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

Author - Aloy Banerjee

Roll No. CH22M503

Q16

```python
# Create a numpy array representing the covariance matrix
covarince_mat = np.array([[1.01, 0.25, -0.02], [0.025, 1.01, 0.04], [-0.017, 0.04, 1.01]])
# Calculate the eigenvalues and corresponding eigenvectors of the covariance matrix
eigenvalues, eigenvectors = np.linalg.eig(covarince_mat)
# Sort the indices of the eigenvalues in descending order
eigenvalue_indices = np.argsort(eigenvalues)[::-1]
# Create a sorted list of eigenvalues
eigenvalues_sorted = eigenvalues[eigenvalue_indices]
# Create a sorted matrix of eigenvectors according to the sorted eigenvalues
eigenvectors_sorted = eigenvectors[:, eigenvalue_indices]
# Create a diagonal matrix of the sorted eigenvalues
Lambda = np.diag(eigenvalues_sorted)
# Create an inverse of the diagonal matrix Lambda by taking the reciprocal of each element on the diagonal
# This is equivalent to creating an inverse of a diagonal matrix, which is simply the reciprocal of
# each diagonal element
Lambda_inv = np.diag(1 / np.diag(Lambda))
# Define a new data point for which we are calculating the Hotelling's T^2 statistic
new_data_point = np.array([0.724, 0.649, 1.857])
# Calculate the centered and scaled data point 'Z' by multiplying the inverse square root of Lambda,
# transposed sorted eigenvectors,
# and transposed new data point
Z = ((np.linalg.inv(np.sqrt(Lambda))).dot(eigenvectors_sorted.T)).dot(new_data_point.T)
# Calculate the Hotelling's T^2 statistic by performing the dot product of transposed Z and Z itself
# This is equivalent to summing the square of each element of Z
T_squared = Z.T @ Z
# Print the computed Hotelling's T^2 statistic
print('T Squared Value: ', T_squared)
```

```
T Squared Value:  1.6237874863769912
```

Time left 0:0

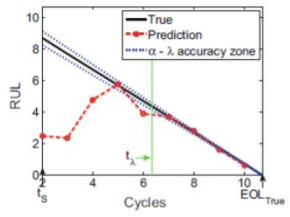
Question **17**

Not yet answered

Marked out of 2.00

Flag question

In an α-λ accuracy analysis for the predicted RUL from the 3rd cycle, λ is considered to be 0.5 and the true end of the life is 10.5 cycles. The relative accuracy is ___ (rounded off to two digits)



Answer: 0.6

ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

Author - Aloy Banerjee

Roll No. CH22M503

Question **18**

Answer saved

Marked out of 2.00

Flag question

Consider two covariance matrices for two classes as follows

$$\begin{bmatrix} 1.01 & 0.25 & -0.02 \\ 0.025 & 1.01 & 0.04 \\ -0.017 & 0.04 & 1.01 \end{bmatrix} \text{ and } \begin{bmatrix} 11 & 2 & -2 \\ 1 & 3 & 4 \\ -1 & 2 & 1 \end{bmatrix}.$$

The similarity index for these two classes for the 2 loading vectors is

Answer: 34.261999058555546

Q18

```python
covarince_mat_1 = np.array([[1.01, 0.25, -0.02], [0.025, 1.01, 0.04], [-0.017, 0.04, 1.01]])
covarince_mat_2 = np.array([[11, 2, -2], [1, 3, 4], [-1, 2, 1]])

def pca_similarity_index(covariance_mat1, covariance_mat2):
    _, eigenvecs1 = np.linalg.eig(covariance_mat1)
    _, eigenvecs2 = np.linalg.eig(covariance_mat2)

    # Sort the eigenvalues and eigenvectors in descending order
    idx1 = np.argsort(eigenvecs1)[::-1]
    eigenvecs1_sorted = eigenvecs1[:, idx1]
    idx2 = np.argsort(eigenvecs2)[::-1]
    eigenvecs2_sorted = eigenvecs2[:, idx2]

    # Compute the PCA similarity index
    similarity_index = np.sum(np.abs(np.dot(eigenvecs1_sorted, eigenvecs2_sorted.T)))

    return similarity_index
```

```python
similarity_index = pca_similarity_index(covarince_mat_1, covarince_mat_2)
print("Similarity Index:", similarity_index)
```

```
Similarity Index: 34.261999058555546
```

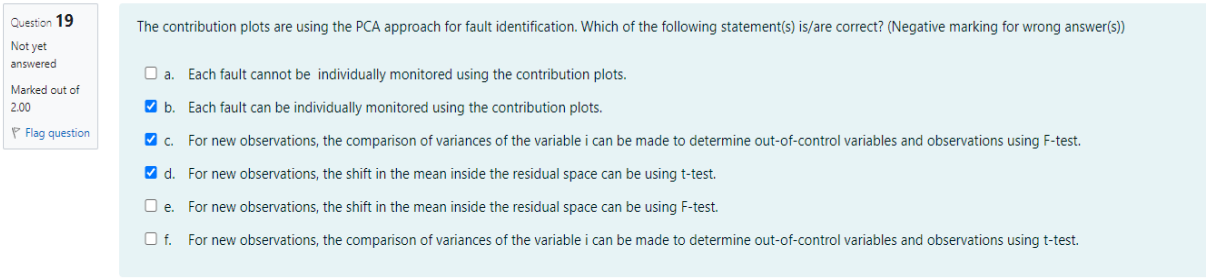
Two vectors if orthogonal to each other then, their orthogonal product must be 0 but if they are in same direction their orthogonal product will not be zero and it will be on higher side. As the value is higher than zero that means they are like each other in some way.

ID5004: AI in Predictive Maintenance, Reliability, and Warranty : End Sem

Part 1

Author - Aloy Banerjee

Roll No. CH22M503



b. Each fault can be individually monitored using the contribution plots.

- Contribution plots in PCA display the relative contribution of variables to each principal component. As each fault will have its unique contribution to the principal components, it becomes possible to monitor each fault individually using these plots.

c. For new observations, the comparison of variances of the variable i can be made to determine out-of-control variables and observations using F-test.

- In PCA-based fault identification, one can use the F-test to compare the variances of variables in new observations. This allows the detection of out-of-control variables and observations within the PCA residual space.

d. For new observations, the shift in the mean inside the residual space can be using t-test.

- This statement is incorrect. The t-test is not typically used to detect a shift in the mean inside the residual space. Instead, the t-test is used to compare the means of two groups, and it's not directly applicable to PCA-based fault identification.

========== Part 1 ==========