Table of Contents

Question 1 (Linear Algebra)

```
if true
    for i =1:5

        fprintf('\nQuestion 1 Variant %d\n',i)
        dim = randi([4 12]);
        % Generated singular values
        singular_values = 100*rand(1,dim);
        % Displaying Question
        disp("Consider a square data matrix M(" + num2str(dim)+ ")."+ "The singular value decomposition (SVD) of M is given by M = UΣV^T, where U is a" );
        disp( num2str(dim) + "x" + num2str(dim) + " orthogonal matrix, Σ is a " + num2str(dim) + "x" + num2str(dim) + " diagonal matrix of singular values, and V^T is the transpose of a
        disp("Singular values of data matrix M [" + num2str(singular_values) + "]");
        disp("Calculate the following:");
        disp("a) The sum of the squared singular values.");
        disp("b) The rank of matrix M.");
        disp("c) The condition number of matrix M.");
        disp(" ");

        % Solution procedure
        disp("Solution:");

        % a) Sum of squared singular values
        sum_squared_singular_values = sum(singular_values.^2);
        disp(['a)  Sum of squared singular values = ' num2str(sum_squared_singular_values)]);

        % b) Rank of matrix M
        rank_M = length(find(singular_values > 0));
        disp(['b)   Rank of matrix M = ' num2str(rank_M)]);

        % c) Condition number of matrix M
        condition_number = max(singular_values) / min(singular_values);
        disp(['c)   Condition number of matrix M = ' num2str(condition_number)]);
        disp(" ");

    end
    disp("--Step by Step Solution--")
    disp("Dimension = "+ num2str(dim()))
    % a) Sum of squared singular values
    disp("a) To calculate the sum of squared singular values, square each singular value and then sum them:");
    disp(['   Sum of squared singular values = ' num2str(singular_values(1)^2) ' + ' num2str(singular_values(2)^2) ' + ' num2str(singular_values(3)^2) ' + ' num2str(singular_values(4)^2
    disp(['   Sum of squared singular values = ' num2str(sum_squared_singular_values)]);
    % b) Rank of matrix M
    rank_M = length(find(singular_values > 0));
    disp("b) To determine the rank of matrix M, count the number of non-zero singular values:");
    disp(['   Rank of matrix M = ' num2str(rank_M)]);
    % c) Condition number of matrix M
```

```
    condition_number = max(singular_values) / min(singular_values);
    disp("c) To calculate the condition number of matrix M, divide the largest singular value by the smallest singular value:");
    disp(['   Condition number of matrix M = ' num2str(max(singular_values)) ' / ' num2str(min(singular_values))]);
    disp(['   Condition number of matrix M = ' num2str(condition_number)]);
    disp(" ")
    % Topic and Subtopic
    disp("Topic: Linear Algebra");
    disp("Subtopic: Singular Value Decomposition (SVD)");
end
```

Question 1 Variant 1
Consider a square data matrix M(12).The singular value decomposition (SVD) of M is given by M = UΣV^T, where U is a
12x12 orthogonal matrix, Σ is a 12x12 diagonal matrix of singular values, and V^T is the transpose of a 12x12 orthogonal matrix.
Singular values of data matrix M [76.464 24.3684 68.2116 13.7855 62.9812 85.7015 89.9798 34.8368 48.631 67.9517 70.4122 46.0884]
Calculate the following:
a) The sum of the squared singular values.
b) The rank of matrix M.
c) The condition number of matrix M.

Solution:
a) Sum of squared singular values = 45969.2095
b) Rank of matrix M = 12
c) Condition number of matrix M = 6.5272

Question 1 Variant 2
Consider a square data matrix M(7).The singular value decomposition (SVD) of M is given by M = UΣV^T, where U is a
7x7 orthogonal matrix, Σ is a 7x7 diagonal matrix of singular values, and V^T is the transpose of a 7x7 orthogonal matrix.
Singular values of data matrix M [28.0265 7.62044 44.4625 16.5706 39.8749 92.0585 51.1334]
Calculate the following:
a) The sum of the squared singular values.
b) The rank of matrix M.
c) The condition number of matrix M.

Solution:
a) Sum of squared singular values = 15774.4539
b) Rank of matrix M = 7
c) Condition number of matrix M = 12.0805

Question 1 Variant 3
Consider a square data matrix M(12).The singular value decomposition (SVD) of M is given by M = UΣV^T, where U is a
12x12 orthogonal matrix, Σ is a 12x12 diagonal matrix of singular values, and V^T is the transpose of a 12x12 orthogonal matrix.
Singular values of data matrix M [9.19339 99.3037 9.64336 31.3147 78.5375 60.24 46.5908 29.8131 13.3173 29.5009 16.6627 31.7104]
Calculate the following:
a) The sum of the squared singular values.
b) The rank of matrix M.
c) The condition number of matrix M.

Solution:
a) Sum of squared singular values = 26206.7327
b) Rank of matrix M = 12
c) Condition number of matrix M = 10.8016

Question 1 Variant 4
Consider a square data matrix M(4).The singular value decomposition (SVD) of M is given by M = UΣV^T, where U is a
4x4 orthogonal matrix, Σ is a 4x4 diagonal matrix of singular values, and V^T is the transpose of a 4x4 orthogonal matrix.
Singular values of data matrix M [83.2088 97.1593 21.8271 70.6078]

Calculate the following:
a) The sum of the squared singular values.
b) The rank of matrix M.
c) The condition number of matrix M.

Solution:
a) Sum of squared singular values = 21825.5082
b)   Rank of matrix M = 4
c)   Condition number of matrix M = 4.4513

Question 1 Variant 5
Consider a square data matrix M(4).The singular value decomposition (SVD) of M is given by M = UΣV^T, where U is a
4x4 orthogonal matrix, Σ is a 4x4 diagonal matrix of singular values, and V^T is the transpose of a 4x4 orthogonal matrix.
Singular values of data matrix M [61.6295 66.9364 3.72015 0.333456]
Calculate the following:
a) The sum of the squared singular values.
b) The rank of matrix M.
c) The condition number of matrix M.

Solution:
a) Sum of squared singular values = 8292.6315
b)   Rank of matrix M = 4
c)   Condition number of matrix M = 200.7353

--Step by Step Solution--
Dimension = 4
a) To calculate the sum of squared singular values, square each singular value and then sum them:
   Sum of squared singular values = 3798.1968 + 4480.4839 + 13.8395 + 0.11119
   Sum of squared singular values = 8292.6315
b) To determine the rank of matrix M, count the number of non-zero singular values:
   Rank of matrix M = 4
c) To calculate the condition number of matrix M, divide the largest singular value by the smallest singular value:
   Condition number of matrix M = 66.9364 / 0.33346
   Condition number of matrix M = 200.7353

Topic: Linear Algebra
Subtopic: Singular Value Decomposition (SVD)

Question 2 (Linear Algebra)

```
if true
% Given vectors
v1 = [1; -1; -2];
v2 = [2; 0; 1];

for i = 1:5
    % Generate random integer input values greater than or equal to zero
    X = randi([0, 10], [3, 1]);

    % Calculate the projection
    projection = ((dot(X, v1) / norm(v1)^2) * v1) + ((dot(X, v2) / norm(v2)^2) * v2);

    % Calculate a*b*c*d
    abcd = projection(1) * projection(2) * projection(3) * dot(X, projection);
```

```
    % Display the modified question and answer

    fprintf("\nQuestion 1 Variant %d\nGiven vector X = [%d; %d; %d]'\n", i, X(1), X(2), X(3));
    fprintf("Projection of X onto the plane defined by v1=[1 -1 -2] and v2=[2 0 1]. X1=[a; b; c] where X1 is Projection of X \n");
    fprintf("and X * X1' = d , where ' represents transpose. Calculate a*b*c*d:\n");
    disp("Options: ");
    fprintf("(A) %f\n", abcd*1.5);
    fprintf("(B) %f\n", abcd*0.33);
    fprintf("(C) %f\n", abcd);
    fprintf("(D) %f\n", abcd*0.5);
    fprintf("Correct Answer %d: %f\n", i, abcd);

    % Print separator
    fprintf(" ");
end

% Explanation :
fprintf("--Step by Step Explanation--\n");
fprintf("Step 1 = to get the projection of X, use the formula to find the projection of X on the plane,")
fprintf("       where v1 and v2 are basis vectors\n");
fprintf("Step 2 = Formula is X1 = (V1'X/V1'V1)V1 + (V2'X/V2'V2)V2\n");
fprintf("Step 3 = find matrix multiplication X.X1\n");
fprintf("Step 4 = From steps 1 and 2, you will get the values of a, b, c, and d, then you can find a*b*c*d");
fprintf(" ");
fprintf("Topic: Linear Algebra");
fprintf("Subtopic: Vector Projection and Dot Product");

end
```

```
Question 1 Variant 1
Given vector X = [1; 9; 3]'
Projection of X onto the plane defined by v1=[1 -1 -2] and v2=[2 0 1]. X1=[a; b; c] where X1 is Projection of X
and X * X1' = d , where ' represents transpose. Calculate a*b*c*d:
Options:
(A) -249.018519
(B) -54.784074
(C) -166.012346
(D) -83.006173
Correct Answer 1: -166.012346

Question 1 Variant 2
Given vector X = [5; 5; 6]'
Projection of X onto the plane defined by v1=[1 -1 -2] and v2=[2 0 1]. X1=[a; b; c] where X1 is Projection of X
and X * X1' = d , where ' represents transpose. Calculate a*b*c*d:
Options:
(A) 7147.008000
(B) 1572.341760
(C) 4764.672000
(D) 2382.336000
Correct Answer 2: 4764.672000

Question 1 Variant 3
Given vector X = [3; 4; 4]'
Projection of X onto the plane defined by v1=[1 -1 -2] and v2=[2 0 1]. X1=[a; b; c] where X1 is Projection of X
and X * X1' = d , where ' represents transpose. Calculate a*b*c*d:
Options:
(A) 942.187500
(B) 207.281250
(C) 628.125000
```

(D) 314.062500
Correct Answer 3: 628.125000

Question 1 Variant 4
Given vector X = [10; 0; 4]'
Projection of X onto the plane defined by v1=[1 -1 -2] and v2=[2 0 1]. X1=[a; b; c] where X1 is Projection of X
and X * X1' = d , where ' represents transpose. Calculate a*b*c*d:
Options:
(A) -2378.613926
(B) -523.295064
(C) -1585.742617
(D) -792.871309
Correct Answer 4: -1585.742617

Question 1 Variant 5
Given vector X = [8; 6; 3]'
Projection of X onto the plane defined by v1=[1 -1 -2] and v2=[2 0 1]. X1=[a; b; c] where X1 is Projection of X
and X * X1' = d , where ' represents transpose. Calculate a*b*c*d:
Options:
(A) 2664.587852
(B) 586.209327
(C) 1776.391901
(D) 888.195951
Correct Answer 5: 1776.391901

--Step by Step Explanation--
Step 1 = to get the projection of X, use the formula to find the projection of X on the plane,
where v1 and v2 are basis vectors
Step 2 = Formula is X1 = (V1'X/V1'V1)V1 + (V2'X/V2'V2)V2
Step 3 = find matrix multiplication X.X1
Step 4 = From steps 1 and 2, you will get the values of a, b, c, and d, then you can find a*b*c*d

Topic: Linear Algebra
Subtopic: Vector Projection and Dot Product

Question 3 (Optimization)

```
if true
 disp('<strong>Question:')
disp('A biotech company is designing a new bioreactor system for producing a protein that is used in medical research. The protein is produced by a culture of bacteria that require cert
disp('The company wants to maximize the yield of the protein by optimizing the concentrations of the nutrients in the bioreactor. The bioreactor has three input streams, each containing
disp('The flow rates of the streams are fixed, but the concentrations of the nutrients can be varied. The concentrations of the nutrients affect the growth rate of the bacteria, as well

disp('The growth rate of the bacteria can be modelled by the function g(a, b, c) = 5a + 3b^2 + 2c, where a, b, and c are the concentrations of the three nutrients, respectively.')

disp('The rate of protein production can be modelled by the function p(a, b, c) = 2a^2 + 3b + 4c^2, where a, b, and c are the concentrations of the three nutrients, respectively.')

disp('The objective function should include both the growth rate and the rate of protein production. Apply the steepest descent optimization method to find the optimal concentrations of
disp('Use an initial point of (a0, b0, c0) = (2, 3, 1) (For different variants different initial points has been taken) and a learning rate of 0.023.')

% Define the growth rate and protein production functions
g = @(a,b,c) 5*a + 3*b^2 + 2*c;
p = @(a,b,c) 2*a^2 + 3*b + 4*c^2;

% Define the objective function
Obj_fun = @(a,b,c) g(a,b,c) * p(a,b,c);
```

```
% Giving negative sign for maximization of yield
% For maximization problems, we typically minimize the negative of the objective function.
% In this case, we want to maximize the yield of the protein, which is determined by the growth rate and the rate of protein production. Therefore, we need to define an objective functi

f= @(a,b,c) -(g(a,b,c) * p(a,b,c));

% Set the initial values for the variant
a = 2;
b_ = 3;
c = 1;

% Set the learning rate
learning_rate = 0.023;

% Define the number of iterations
num_iterations = 2;

% Define the number of variants
num_variants = 6;

% Define the initial perturbation
perturbation = 0;

% Perform the steepest descent optimization for each variant


fprintf('\n')


for v = 1:num_variants
    % Set the initial values for the perturbed variant
    %The variants are created by addig 0.2 to the initial points and so on
    %Here we have taken 6 variants
    a_perturbed = a + perturbation;
    b_perturbed = b_ + perturbation;
    c_perturbed = c + perturbation;

    fprintf('\n')
    % Display the initial values for the perturbed variant
    disp(['<strong>Variant ', num2str(v), ' initial points are:(a,b,c)=(', num2str(a_perturbed), ',', num2str(b_perturbed), ',', num2str(c_perturbed), ')'])
    fprintf('\n')

    % Update the perturbation for the next variant
    perturbation = perturbation + 0.2;

    % Perform the steepest descent optimization
    for i = 1:num_iterations
        % Calculate the partial derivatives of f with respect to a, b, and c
        df_da = -30*a_perturbed^2 - 12*a_perturbed*b_perturbed^2 - 8*a_perturbed*c_perturbed - 15*b_perturbed - 20*c_perturbed^2;
        df_db = -12*a_perturbed^2*b_perturbed - 15*a_perturbed - 27*b_perturbed^2 - 24*b_perturbed*c_perturbed^2 - 6*c_perturbed;
        df_dc = -4*a_perturbed^2 - 40*a_perturbed*c_perturbed - 24*b_perturbed^2*c_perturbed - 6*b_perturbed - 24*c_perturbed^2;

        % Update the values of a, b, and c
        a_perturbed = a_perturbed - learning_rate * df_da;
        b_perturbed = b_perturbed - learning_rate * df_db;
        c_perturbed = c_perturbed - learning_rate * df_dc;

        % Calculate the value of the objective function
        obj = Obj_fun(a_perturbed,b_perturbed,c_perturbed);
        % Display the results
```

```
        disp(['Variant ', num2str(v), ', Iteration ', num2str(i)])

        disp(['a = ', num2str(a_perturbed)])
        disp(['b = ', num2str(b_perturbed)])
        disp(['c = ', num2str(c_perturbed)])
        disp(['Objective function = ', num2str(obj)])
    end
end



%%
fprintf('\n')
disp('<strong>Explanation for answer:')
fprintf('\n')
disp('The yield of protein refers to the amount of protein produced by a culture of bacteria under specific conditions. ')
disp('The growth rate of the bacteria determines the number of bacteria present in the culture, while the rate of protein production determines the amount of protein produced per unit c
disp('The yield of protein is thus the product of the total amount of protein produced and the number of bacteria present in the culture.')
disp('If we assume that the protein production rate is proportional to the number of bacteria present, then we can write the yield of protein as:')
fprintf('\n')
disp('Yield = Protein production rate * Number of bacteria')
fprintf('\n')
disp('In turn, the number of bacteria is determined by the growth rate. The higher the growth rate, the faster the bacteria divide and the more bacteria are present in the culture. ')
disp('Therefore, we can write the number of bacteria as a function of the nutrient concentrations, which is given by the growth rate function g(a, b, c).')
disp('Finally, the protein production rate is also a function of the nutrient concentrations, which is given by the protein production rate function p(a, b, c). ')
disp('By multiplying these two functions together, we obtain the objective function f(a, b, c, d) that represents the yield of protein in terms of the nutrient concentrations.')

end
```

Question:

A biotech company is designing a new bioreactor system for producing a protein that is used in medical research. The protein is produced by a culture of bacteria that require certain nutrient

The company wants to maximize the yield of the protein by optimizing the concentrations of the nutrients in the bioreactor. The bioreactor has three input streams, each containing a different

The flow rates of the streams are fixed, but the concentrations of the nutrients can be varied. The concentrations of the nutrients affect the growth rate of the bacteria, as well as the rate

The growth rate of the bacteria can be modelled by the function g(a, b, c) = 5a + 3b^2 + 2c, where a, b, and c are the concentrations of the three nutrients, respectively.

The rate of protein production can be modelled by the function p(a, b, c) = 2a^2 + 3b + 4c^2, where a, b, and c are the concentrations of the three nutrients, respectively.

The objective function should include both the growth rate and the rate of protein production. Apply the steepest descent optimization method to find the optimal concentrations of the three r

Use an initial point of (a0, b0, c0) = (2, 3, 1) (For different variants different initial points has been taken) and a learning rate of 0.023.

Variant 1 initial points are:(a,b,c)=(2,3,1)

Variant 1, Iteration 1

a = 11.591

b = 14.385

c = 9.142

Objective function = 450390.3638

Variant 1, Iteration 2

a = 829.1862

b = 1345.1956

c = 1211.3485

Objective function = 39397731382166.24

Variant 2 initial points are:(a,b,c)=(2.2,3.2,1.2)

Variant 2, Iteration 1

a = 14.0095

b = 17.3019

c = 12.0935

Objective function = 1021531.5096

Variant 2, Iteration 2
a = 1411.3532
b = 2543.7605
c = 2267.5389
Objective function = 477016002901225.6
Variant 3 initial points are:(a,b,c)=(2.4,3.4,1.4)
Variant 3, Iteration 1
a = 16.7246
b = 20.6837
c = 15.5058
Objective function = 2213428.8149
Variant 3, Iteration 2
a = 2349.9602
b = 4636.1378
c = 4077.1436
Objective function = 5002130260509768
Variant 4 initial points are:(a,b,c)=(2.6,3.6,1.6)
Variant 4, Iteration 1
a = 19.7495
b = 24.5699
c = 19.4053
Objective function = 4598823.7465
Variant 4, Iteration 2
a = 3831.6841
b = 8161.1701
c = 7085.6039
Objective function = 4.60071185869247e+16
Variant 5 initial points are:(a,b,c)=(2.8,3.8,1.8)
Variant 5, Iteration 1
a = 23.0976
b = 29.0005
c = 23.8185
Objective function = 9195659.4115
Variant 5, Iteration 2
a = 6124.9025
b = 13914.5747
c = 11953.8818
Objective function = 3.756404629883749e+17
Variant 6 initial points are:(a,b,c)=(3,4,2)
Variant 6, Iteration 1
a = 26.782
b = 34.015
c = 28.772
Objective function = 17755534.1344
Variant 6, Iteration 2
a = 9608.5225
b = 23043.1479
c = 19641.3043
Objective function = 2.752530684950067e+18
Explanation for answer:
The yield of protein refers to the amount of protein produced by a culture of bacteria under specific conditions.
The growth rate of the bacteria determines the number of bacteria present in the culture, while the rate of protein production determines the amount of protein produced per unit of time.
The yield of protein is thus the product of the total amount of protein produced and the number of bacteria present in the culture.
If we assume that the protein production rate is proportional to the number of bacteria present, then we can write the yield of protein as:
Yield = Protein production rate * Number of bacteria

In turn, the number of bacteria is determined by the growth rate. The higher the growth rate, the faster the bacteria divide and the more bacteria are present in the culture.
Therefore, we can write the number of bacteria as a function of the nutrient concentrations, which is given by the growth rate function g(a, b, c).
Finally, the protein production rate is also a function of the nutrient concentrations, which is given by the protein production rate function p(a, b, c).
By multiplying these two functions together, we obtain the objective function f(a, b, c, d) that represents the yield of protein in terms of the nutrient concentrations.

Question 4 (Optimization)

```
disp("Question 4 will be here");
```

```
Question 4 will be here
```

Question 5 (Statistics & Probability)

```
if true
    data_point = 5;
    for i = 1:5
        fprintf('<strong>\nQuestion 5 Variant %d\n <strong>',i)
        data_point = randi([3 10]);

        X1 = linspace(1,data_point,data_point);
        X2 = 2*X1;
        X3 = 3*X1;
        Y = randi(100,1,data_point);

        disp("You are working on a machine learning project and have collected a dataset with " + num2str(data_point) +" samples. The dataset ");
        disp("contains three features: X1, X2, and X3. You want to determine the relationship between these features ");
        disp("and a target variable Y. After performing linear regression, you obtain the following coefficient estimates:");
        beta0 = 4.5;
        disp("β0 (intercept) = " + num2str(beta0));
        beta1 = 2.8;
        disp("β1 (coefficient for X1) = " + num2str(beta1));
        beta2 = -1.5;
        disp("β2 (coefficient for X2) = " + num2str(beta2));
        beta3 = 3.5;
        disp("β3 (coefficient for X3) = " + num2str(beta3));
        disp("");
        disp("Sample data");
        disp("X1 = ["+ num2str(X1) + "]");
        disp("X2 = ["+ num2str(X2) + "]");
        disp("X3 = ["+ num2str(X3) + "]");
        disp("Y = ["+ num2str(Y) + "]");
        disp("Please provide your answers to following questions: ");
        disp("a) Write the equation of the linear regression model.");
        disp("b) What does the coefficient β1 represent in the linear regression model?");
        disp("c) Calculate the coefficient of determination (R-squared).");
        % Calculate the predicted values Ŷi

        Y_predicted = beta0 + beta1*X1 + beta2*X2 + beta3*X3;

        % Calculate the mean value of Y (Ȳ)
        Y_mean = mean(Y);

        % Calculate the coefficient of determination (R-squared)
```

```
    SSR = sum((Y_predicted - Y_mean).^2); % sum of square residual
    SST = sum((Y - Y_mean).^2); % sum of squares total
    R_squared = (SSR / SST); % coeffient of determination

    disp(" ");
    % Display the results
    disp("<strong>-Answer-<strong>");
    disp("a) Equation of the linear regression model: Y = " +  num2str(beta0)+" + " + num2str(beta1) + "*X1 " + num2str(beta2)+"*X2 + " + num2str(beta3)+"*X3");
    disp("b) Coefficient β1 (" + num2str(beta1)+ ")" + " represents the change in the target variable Y for a one-unit increase in X1,");
    disp(" while holding other variables constant.");
    if R_squared <1
        disp("c) Coefficient of determination (R-squared): " + num2str(R_squared));
    else
        disp("c) Data used to calculate the R-squared value is invalid. Since values of R-squared = " + num2str(R_squared) + " (>1).")
        disp(" which is outside the recommended range");
    end
    disp (" ");
end
% Explanation
disp("<strong>--Step by Step Solution--<strong>")
disp("No. of data point = "+ num2str(data_point));
% Problem 1 explanation
disp("<strong>a) Equation of the linear regression model.<strong>");
disp("The equation of the linear regression model is given by: ");
disp("Y = β0 + β1*X1 + β2*X2 + β3*X3");
disp("Substituting the coefficient values, we have: ");
disp("Y = " +  num2str(beta0)+" + " + num2str(beta1) + "*X1 " + num2str(beta2)+"*X2 + " + num2str(beta3)+"*X3");
disp(" ");
% Problem 2 explanation
disp("<strong>b) Interpret the coefficient β1 in the context of the problem. <strong>");
disp("The coefficient β1 represents the effect of the feature X1 on the target variable Y. In this context, for every unit increase in X1, ");
disp("the target variable Y is expected to increase by " + num2str(beta1) + " units, assuming all other variables remain constant.");
disp(" ");
% Problem 3 explanation
disp("<strong>c) Determine the coefficient of determination (R-squared) <strong>");
disp("To calculate the coefficient of determination (R-squared), we need to compute the total sum of squares (SST) and the sum of squared errors (SSE).");
disp("<strong>Step 1: <strong> Calculate the mean of the target variable Y:");
disp("Y_mean = (ΣY / N), where N is the number of samples");
disp(" ");
disp("<strong>Step 2: <strong> Calculate the total sum of squares (SST):");
disp("SST = Σ((Y - Y_mean)^2) = " + num2str(SST));
disp(" ");
disp("<strong>Step 3: <strong> Calculate the predicted values of Y (Y_predicted) using the linear regression model:");
disp("Y_predicted = "  + num2str(beta0) + " + "+ num2str(beta1)+ "*X1 " + num2str(beta2) + "*X2 + " + num2str(beta3) + "*X3");
disp(" ");
disp("<strong>Step 4: <strong> Calculate the sum of squared errors (SSE):");
disp("SSR = Σ((Y_mean - Y_predicted)^2) = " + num2str(SSR));
disp(" ");
disp("<strong>Step 5: <strong> Calculate the coefficient of determination (R-squared):");
disp("R_squared = SSR / SST = " +  num2str(SSR/SST) + " = " +num2str(R_squared));
disp(" ");
disp("<strong>Note: <strong> R-squared measures the proportion of the total variation in the target variable that is explained by the linear regression model.");
disp("It ranges from 0 to 1, where 1 indicates perfect fit and 0 indicates no fit.");
disp("R-squared values outside of the 0-1 range are not meaningful or useful for interpreting the goodness-of-fit of a model. ");
disp("It could be an indication that the data used to calculate the R-squared value is invalid.");
disp(" ");
disp("Therefore, it is important to ensure that the calculation and data used to calculate R-squared are ");
disp("correct and valid, and if the value still falls outside of the range, it may be necessary to review ");
disp("the underlying assumptions and methodology used in the analysis.");
disp(" ");
disp("<strong>Topic: <strong> Statistics" );
```

```
    subtopics = {'Linear regression', 'Coefficient interpretation', 'Coefficient of determination (R-squared)'};

    disp("<strong>Subtopics: <strong>");
    for i = 1:length(subtopics)
        disp("- " + num2str(subtopics{i}));
    end

end
```

Question 5 Variant 1

You are working on a machine learning project and have collected a dataset with 10 samples. The dataset contains three features: X1, X2, and X3. You want to determine the relationship between these features and a target variable Y. After performing linear regression, you obtain the following coefficient estimates:
β0 (intercept) = 4.5
β1 (coefficient for X1) = 2.8
β2 (coefficient for X2) = -1.5
β3 (coefficient for X3) = 3.5
Sample data
X1 = [1 2 3 4 5 6 7 8 9 10]
X2 = [2 4 6 8 10 12 14 16 18 20]
X3 = [3 6 9 12 15 18 21 24 27 30]
Y = [10 32 89 66 69 48 15 96 89 44]
Please provide your answers to following questions:
a) Write the equation of the linear regression model.
b) What does the coefficient β1 represent in the linear regression model?
c) Calculate the coefficient of determination (R-squared).

-Answer-
a) Equation of the linear regression model: Y = 4.5 + 2.8*X1 -1.5*X2 + 3.5*X3
b) Coefficient β1 (2.8) represents the change in the target variable Y for a one-unit increase in X1, while holding other variables constant.
c) Data used to calculate the R-squared value is invalid. Since values of R-squared = 1.0476 (>1). which is outside the recommended range

Question 5 Variant 2

You are working on a machine learning project and have collected a dataset with 9 samples. The dataset contains three features: X1, X2, and X3. You want to determine the relationship between these features and a target variable Y. After performing linear regression, you obtain the following coefficient estimates:
β0 (intercept) = 4.5
β1 (coefficient for X1) = 2.8
β2 (coefficient for X2) = -1.5
β3 (coefficient for X3) = 3.5
Sample data
X1 = [1 2 3 4 5 6 7 8 9]
X2 = [2 4 6 8 10 12 14 16 18]
X3 = [3 6 9 12 15 18 21 24 27]
Y = [33 37 80 10 96 1 30 5 45]
Please provide your answers to following questions:
a) Write the equation of the linear regression model.
b) What does the coefficient β1 represent in the linear regression model?
c) Calculate the coefficient of determination (R-squared).

-Answer-

a) Equation of the linear regression model: Y = 4.5 + 2.8*X1 -1.5*X2 + 3.5*X3
b) Coefficient β1 (2.8) represents the change in the target variable Y for a one-unit increase in X1,
while holding other variables constant.
c) Data used to calculate the R-squared value is invalid. Since values of R-squared = 1.1126 (>1).
which is outside the recommended range

Question 5 Variant 3

You are working on a machine learning project and have collected a dataset with 9 samples. The dataset contains three features: X1, X2, and X3. You want to determine the relationship between these features and a target variable Y. After performing linear regression, you obtain the following coefficient estimates:
β0 (intercept) = 4.5
β1 (coefficient for X1) = 2.8
β2 (coefficient for X2) = -1.5
β3 (coefficient for X3) = 3.5
Sample data
X1 = [1 2 3 4 5 6 7 8 9]
X2 = [2 4 6 8 10 12 14 16 18]
X3 = [3 6 9 12 15 18 21 24 27]
Y = [92 54 81 57 76 1 48 26 31]
Please provide your answers to following questions:
a) Write the equation of the linear regression model.
b) What does the coefficient β1 represent in the linear regression model?
c) Calculate the coefficient of determination (R-squared).

-Answer-
a) Equation of the linear regression model: Y = 4.5 + 2.8*X1 -1.5*X2 + 3.5*X3
b) Coefficient β1 (2.8) represents the change in the target variable Y for a one-unit increase in X1,
while holding other variables constant.
c) Coefficient of determination (R-squared): 0.96258

Question 5 Variant 4

You are working on a machine learning project and have collected a dataset with 10 samples. The dataset contains three features: X1, X2, and X3. You want to determine the relationship between these features and a target variable Y. After performing linear regression, you obtain the following coefficient estimates:
β0 (intercept) = 4.5
β1 (coefficient for X1) = 2.8
β2 (coefficient for X2) = -1.5
β3 (coefficient for X3) = 3.5
Sample data
X1 = [1 2 3 4 5 6 7 8 9 10]
X2 = [2 4 6 8 10 12 14 16 18 20]
X3 = [3 6 9 12 15 18 21 24 27 30]
Y = [21 53 23 57 100 14 96 13 19 65]
Please provide your answers to following questions:
a) Write the equation of the linear regression model.
b) What does the coefficient β1 represent in the linear regression model?
c) Calculate the coefficient of determination (R-squared).

-Answer-
a) Equation of the linear regression model: Y = 4.5 + 2.8*X1 -1.5*X2 + 3.5*X3
b) Coefficient β1 (2.8) represents the change in the target variable Y for a one-unit increase in X1,
while holding other variables constant.

c) Data used to calculate the R-squared value is invalid. Since values of R-squared = 1.1081 (>1).
which is outside the recommended range

Question 5 Variant 5

You are working on a machine learning project and have collected a dataset with 4 samples. The dataset contains three features: X1, X2, and X3. You want to determine the relationship between these features and a target variable Y. After performing linear regression, you obtain the following coefficient estimates:
β0 (intercept) = 4.5
β1 (coefficient for X1) = 2.8
β2 (coefficient for X2) = -1.5
β3 (coefficient for X3) = 3.5
Sample data
X1 = [1 2 3 4]
X2 = [2 4 6 8]
X3 = [3 6 9 12]
Y = [9 66 3 99]
Please provide your answers to following questions:
a) Write the equation of the linear regression model.
b) What does the coefficient β1 represent in the linear regression model?
c) Calculate the coefficient of determination (R-squared).

-Answer-
a) Equation of the linear regression model: Y = 4.5 + 2.8*X1 -1.5*X2 + 3.5*X3
b) Coefficient β1 (2.8) represents the change in the target variable Y for a one-unit increase in X1, while holding other variables constant.
c) Coefficient of determination (R-squared): 0.20491

--Step by Step Solution--
No. of data point = 4
a) Equation of the linear regression model.
The equation of the linear regression model is given by:
Y = β0 + β1*X1 + β2*X2 + β3*X3
Substituting the coefficient values, we have:
Y = 4.5 + 2.8*X1 -1.5*X2 + 3.5*X3

b) Interpret the coefficient β1 in the context of the problem.
The coefficient β1 represents the effect of the feature X1 on the target variable Y. In this context, for every unit increase in X1, the target variable Y is expected to increase by 2.8 units, assuming all other variables remain constant.

c) Determine the coefficient of determination (R-squared)
To calculate the coefficient of determination (R-squared), we need to compute the total sum of squares (SST) and the sum of squared errors (SSE).
Step 1: Calculate the mean of the target variable Y:
Y_mean = (ΣY / N), where N is the number of samples

Step 2: Calculate the total sum of squares (SST):
SST = Σ((Y - Y_mean)^2) = 6414.75

Step 3: Calculate the predicted values of Y (Y_predicted) using the linear regression model:
Y_predicted = 4.5 + 2.8*X1 -1.5*X2 + 3.5*X3

Step 4: Calculate the sum of squared errors (SSE):
SSR = Σ((Y_mean - Y_predicted)^2) = 1314.45

Step 5: Calculate the coefficient of determination (R-squared):
R_squared = SSR / SST = 0.20491 = 0.20491

Note: R-squared measures the proportion of the total variation in the target variable that is explained by the linear regression model.
It ranges from 0 to 1, where 1 indicates perfect fit and 0 indicates no fit.
R-squared values outside of the 0-1 range are not meaningful or useful for interpreting the goodness-of-fit of a model.
It could be an indication that the data used to calculate the R-squared value is invalid.

Therefore, it is important to ensure that the calculation and data used to calculate R-squared are
correct and valid, and if the value still falls outside of the range, it may be necessary to review
the underlying assumptions and methodology used in the analysis.

Topic: Statistics
Subtopics:
- Linear regression
- Coefficient interpretation
- Coefficient of determination (R-squared)

Question 6 (Statistics & Probability)

```
if true
for i = 1:5
    % Generate random input values while satisfying the condition
    total_coins = randi([4, 10]);
    biased_coins = randi([1, total_coins-1]);
    unbiased_coins = total_coins - biased_coins;

    % Calculate the probability of getting heads on both tosses
    p_heads = 1/4;

    % Calculate the probability of drawing a biased coin
    p_biased = biased_coins / total_coins;

    % Calculate the probability of getting two heads with an unbiased coin
    p_two_heads_unbiased = (1/2)^2;

    % Calculate the probability of getting two heads with a biased coin
    p_two_heads_biased = 1;

    % Calculate the probability of getting two heads
    p_two_heads = (unbiased_coins/total_coins * p_two_heads_unbiased) + (p_biased * p_two_heads_biased);

    % Calculate the probability that the coin is biased given that it gave two heads
    p_biased_given_two_heads = (p_biased * p_two_heads_biased) / p_two_heads;

    % Display the question and answer
    fprintf("<strong>Question 6 Variant %d: There are %d coins, of these %d are unbiased and %d is a biased coin with two heads.\n", i, total_coins, unbiased_coins, biased_coins);
    fprintf("            A coin is drawn at random and tossed two times, it appears head on both the times.\n");
    fprintf("            What is the probability that the head is happened in the biased coin?\n");
    fprintf("Options: (A) %0.2f\n", p_biased_given_two_heads*0.5);
    fprintf("         (B) %0.2f\n", p_biased_given_two_heads*2);
    fprintf("         (C) %0.2f\n", p_biased_given_two_heads);
    fprintf("         (D) %0.2f\n", p_biased_given_two_heads*1.5);
    fprintf("<strong>Answer <strong> %d: %0.2f\n", i, p_biased_given_two_heads);

    % Print separator
    fprintf("--------------------------------------------------\n");
end
```

```
% SOLUTION :
fprintf("<strong>--Step by Step Solution <strong>:\n");
fprintf("Probability of getting head on unbiased coin is 1/2 , i.e P(H) = 1/2\n");
fprintf("Probability of getting head on biased coin is 1, i.e P(H) = 1\n");
fprintf("Let B represent biased and U represent unbiased coins\n");
fprintf("P(B/H) = number of biased coins/total number of coins\n");
fprintf("P(U/H) = 0.5 * (number of unbiased coins/total number of coins)\n");
fprintf("Use the formula to get the answer:\n");
fprintf("P(H/B) = P(H) * P(B/H) / (P(H) * P(B/H) + P(H) * P(U/H))");
disp(" ");
disp("<strong>Topic: <strong> Probability")
disp("<strong>Subtopic:<strong> Coin tossing")

end
```

Question 6 Variant 1: There are 7 coins, of these 4 are unbiased and 3 is a biased coin with two heads.
A coin is drawn at random and tossed two times, it appears head on both the times.
What is the probability that the head is happened in the biased coin?

Options: (A) 0.38
(B) 1.50
(C) 0.75
(D) 1.12

Answer 1: 0.75

Question 6 Variant 2: There are 8 coins, of these 1 are unbiased and 7 is a biased coin with two heads.
A coin is drawn at random and tossed two times, it appears head on both the times.
What is the probability that the head is happened in the biased coin?

Options: (A) 0.48
(B) 1.93
(C) 0.97
(D) 1.45

Answer 2: 0.97

Question 6 Variant 3: There are 6 coins, of these 2 are unbiased and 4 is a biased coin with two heads.
A coin is drawn at random and tossed two times, it appears head on both the times.
What is the probability that the head is happened in the biased coin?

Options: (A) 0.44
(B) 1.78
(C) 0.89
(D) 1.33

Answer 3: 0.89

Question 6 Variant 4: There are 8 coins, of these 2 are unbiased and 6 is a biased coin with two heads.
A coin is drawn at random and tossed two times, it appears head on both the times.
What is the probability that the head is happened in the biased coin?

Options: (A) 0.46
(B) 1.85
(C) 0.92
(D) 1.38

Answer 4: 0.92

Question 6 Variant 5: There are 8 coins, of these 5 are unbiased and 3 is a biased coin with two heads.
A coin is drawn at random and tossed two times, it appears head on both the times.
What is the probability that the head is happened in the biased coin?

Options: (A) 0.35
(B) 1.41

(C) 0.71

(D) 1.06

Answer 5: 0.71

--Step by Step Solution :

Probability of getting head on unbiased coin is 1/2 , i.e P(H) = 1/2

Probability of getting head on biased coin is 1, i.e P(H) = 1

Let B represent biased and U represent unbiased coins

P(B/H) = number of biased coins/total number of coins

P(U/H) = 0.5 * (number of unbiased coins/total number of coins)

Use the formula to get the answer:

P(H/B) = P(H) * P(B/H) / (P(H) * P(B/H) + P(H) * P(U/H))

Topic: Probability

Subtopic: Coin tossing