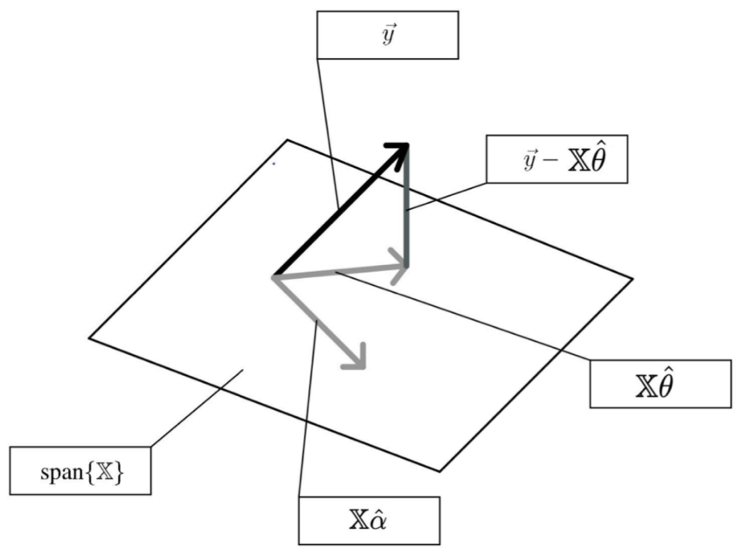
1.



Solution:

(a) What is always true about the residuals in least squares regression? Select all that apply.

- ☐ A. They are orthogonal to the column space of the design matrix.
- ☐ B. They represent the errors of the predictions.
- ☐ C. Their sum is equal to the mean squared error.
- ☐ D. Their sum is equal to zero.
- ☐ E. None of the above.

Solution: (A), (B)

(C): (C) is wrong because the mean squared error is the *mean* of the sum of the *squares* of the residuals.

(D): A counter-example is: $\mathbb{X} = \begin{bmatrix} 2 & 3 \\ 1 & 5 \\ 2 & 4 \end{bmatrix}, \mathbb{Y} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$. After solving the least squares problem, the sum of the residuals is -0.0247, which is not equal to zero. However, note that this statement is in general true if every feature contains the same constant intercept term.

(E): is wrong since (A) and (B) are correct.

(b)

(a) What is always true about the residuals in least squares regression? Select all that apply.

☐ A. They are orthogonal to the column space of the design matrix.
☐ B. They represent the errors of the predictions.
☐ C. Their sum is equal to the mean squared error.
☐ D. Their sum is equal to zero.
☐ E. None of the above.

Solution: (A), (B)

(C): (C) is wrong because the mean squared error is the *mean* of the sum of the *squares* of the residuals.

(D): A counter-example is: $\mathbb{X} = \begin{bmatrix} 2 & 3 \\ 1 & 5 \\ 2 & 4 \end{bmatrix}, \mathbb{Y} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$. After solving the least squares problem, the sum of the residuals is -0.0247, which is not equal to zero. However, note that this statement is in general true if every feature contains the same constant intercept term.

(E): is wrong since (A) and (B) are correct.

(c) We fit a simple linear regression to our data $(x_i, y_i), i = 1, 2, 3$, where x_i is the independent variable and y_i is the dependent variable. Our regression line is of the form $\hat{y} = \hat{\theta}_0 + \hat{\theta}_1 x$. Suppose we plot the relationship between the residuals of the model and the $\hat{y}s$, and find that there is a curve. What does this tell us about our model?

☐ A. The relationship between our dependent and independent variables is well represented by a line.
☐ B. The accuracy of the regression line varies with the size of the dependent variable.
☐ C. The variables need to be transformed, or additional independent variables are needed.

Solution:
If we see a curve in our residual plot, then the relationship is not well represented by a line. Either more independent variables are needed, or transformations of the current variables are necessary.

2.

(a) What is always true about the residuals in least squares regression? Select all that apply.

- ☐ A. They are orthogonal to the column space of the design matrix.
- ☐ B. They represent the errors of the predictions.
- ☐ C. Their sum is equal to the mean squared error.
- ☐ D. Their sum is equal to zero.
- ☐ E. None of the above.

Solution: (A), (B)

(C): (C) is wrong because the mean squared error is the *mean* of the sum of the *squares* of the residuals.

(D): A counter-example is: $\mathbb{X} = \begin{bmatrix} 2 & 3 \\ 1 & 5 \\ 2 & 4 \end{bmatrix}$, $\mathbb{Y} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$. After solving the least squares problem, the sum of the residuals is -0.0247, which is not equal to zero. However, note that this statement is in general true if every feature contains the same constant intercept term.

(E): is wrong since (A) and (B) are correct.