Q1. Simple Linear Regression vs Multiple Linear Regression

Aspect	Simple Linear Regression	Multiple Linear Regression
Definition	Predicts a dependent variable (Y) from **one independent variable (X)**	Predicts a dependent variable (Y) from **two or more independent variables (X1, X2, …)**
Equation	Y=β0+β1X+ϵY = \beta_0 + \beta_1 X + \epsilonY=β0+β1X+ϵ	Y=β0+β1X1+β2X2+…+βnXn+ϵY = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + … + \beta_n X_n + \epsilonY=β0+β1X1+β2X2+…+βnXn+ϵ
Use case	Predicting house price based on square footage	Predicting house price based on square footage, number of bedrooms, age of house, and location

Example:

- Simple: Predict **salary** based on **years of experience**
- Multiple: Predict **salary** based on **years of experience, education level, and location**

Q2. Assumptions of Linear Regression

1. **Linearity**: Relationship between independent and dependent variables is linear.
    - **Check:** Scatter plots, residual plots
2. **Independence of errors**: Residuals are independent.
    - **Check:** Durbin-Watson test
3. **Homoscedasticity**: Constant variance of errors.
    - **Check:** Residual vs predicted value plot
4. **Normality of errors**: Residuals are normally distributed.

- **Check:** Histogram or Q-Q plot of residuals

5. **No multicollinearity (for multiple regression)**: Independent variables are not highly correlated.

    - **Check:** Correlation matrix, VIF (Variance Inflation Factor)

Q3. Interpreting Slope and Intercept

- **Intercept (β_0)**: Predicted value of Y when all X = 0
- **Slope (β_1)**: Change in Y for a 1-unit increase in X

Example:

- Model: $\text{Salary} = 30000 + 5000 \times \text{Years of Experience}$
    - Intercept = 30000 → Base salary with 0 years experience
    - Slope = 5000 → Each additional year of experience increases salary by 5000

Q4. Gradient Descent

- **Concept:** Iterative optimization algorithm to **minimize the loss function** (e.g., Mean Squared Error in regression)
- **How it works:**
    1. Start with random coefficients (β_0, β_1)
    2. Compute gradient (slope) of the loss function with respect to coefficients
    3. Update coefficients in the opposite direction of the gradient
       $\beta = \beta - \alpha \frac{\partial J}{\partial \beta}$

4. Repeat until convergence

- **Use in ML:** Finds optimal parameters for regression, logistic regression, and neural networks

---

Q5. Multiple Linear Regression

- **Model:** Predicts Y using multiple independent variables.
  Y=β0+β1X1+β2X2+…+βnXn+ϵY = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + … + \beta_n X_n + \epsilonY=β0+β1X1+β2X2+…+βnXn+ϵ

- **Difference from Simple Regression:** Multiple predictors instead of one; allows capturing more complex relationships.

Example: Predicting house price using:

- Square footage
- Number of bedrooms
- Age of house
- Location rating

Q6. Multicollinearity in Multiple Linear Regression

- **Definition:** High correlation between two or more independent variables
- **Problem:** Inflates standard errors, unstable coefficients, reduces interpretability
- **Detection:**
  - Correlation matrix (look for r > 0.8)
  - Variance Inflation Factor (VIF > 10 indicates serious multicollinearity)

- **Solutions:**
    - Remove correlated variables
    - Combine features using PCA
    - Regularization (Ridge or Lasso regression)

Q7. Polynomial Regression

- **Concept:** Extends linear regression by including **polynomial terms of the independent variable(s)**
- **Equation:**
 Y=β0+β1X+β2X2+…+βnXn+ϵY = \beta_0 + \beta_1 X + \beta_2 X^2 + … + \beta_n X^n + \epsilonY=β0+β1X+β2X2+…+βnXn+ϵ
- **Difference from linear regression:** Captures **non-linear relationships**

Example: Predicting **growth of bacteria** over time: Y=β0+β1t+β2t2Y = \beta_0 + \beta_1 t + \beta_2 t^2Y=β0+β1t+β2t2

Q8. Advantages and Disadvantages of Polynomial Regression

Advantages:

- Captures non-linear relationships
- Flexible in fitting curved data

Disadvantages:

- Prone to overfitting if degree is high

- Coefficients may lose interpretability
- Sensitive to outliers

When to use:

- Use when **scatter plots show a curved pattern** between X and Y
- Example: Predicting **temperature variation over the day**
