**Q1. What is Lasso Regression, and how does it differ from other regression techniques?**

**Lasso Regression (Least Absolute Shrinkage and Selection Operator)** is a type of linear regression that includes a regularization term. It applies L1 regularization, which adds a penalty equal to the absolute value of the coefficients, encouraging sparsity in the model (i.e., forcing some coefficients to be exactly zero). This makes Lasso useful for feature selection, as it can eliminate irrelevant features by reducing their coefficients to zero.

**Difference from other regression techniques**:

* **Linear Regression**: Lasso differs from simple linear regression by including a penalty term to prevent overfitting and improve generalization.
* **Ridge Regression**: Ridge uses L2 regularization (penalty on squared coefficients), which doesn’t lead to sparse coefficients, unlike Lasso that can set some coefficients to zero.
* **ElasticNet**: Combines both L1 and L2 regularization (a mix of Ridge and Lasso), offering flexibility in feature selection and regularization.

**Q2. What is the main advantage of using Lasso Regression in feature selection?**

The **main advantage** of Lasso Regression in feature selection is its ability to **shrink some coefficients to exactly zero**, effectively removing certain features from the model. This makes Lasso a powerful tool for **automatic feature selection** and for handling high-dimensional datasets, where many features may be irrelevant or redundant.

**Q3. How do you interpret the coefficients of a Lasso Regression model?**

The coefficients in a Lasso regression model can be interpreted similarly to those in linear regression. A positive coefficient indicates a positive relationship with the target variable, and a negative coefficient indicates a negative relationship. The magnitude of the coefficient represents the strength of the relationship. However, due to Lasso's regularization, some coefficients may be zero, indicating that those features are excluded from the model.

**Q4. What are the tuning parameters that can be adjusted in Lasso Regression, and how do they affect the model's performance?**

The key tuning parameter in Lasso Regression is the **regularization parameter (lambda or alpha)**. It controls the strength of the penalty term:

* **Lambda (α)**: When α is zero, Lasso becomes equivalent to linear regression (no regularization). As α increases, the penalty on the coefficients becomes stronger, and more coefficients are shrunk towards zero, increasing sparsity. An excessively high α value may result in underfitting.
* **Optimal lambda**: The optimal value is typically determined using techniques like cross-validation to balance the trade-off between model complexity and prediction accuracy.

**Q5. Can Lasso Regression be used for non-linear regression problems? If yes, how?**

Lasso Regression, being a linear model, is inherently not suited for non-linear relationships. However, it can be used for **non-linear regression problems** by **transforming the features** into a higher-dimensional space (for example, using polynomial or interaction features) and then applying Lasso. This approach allows Lasso to handle non-linearities in the transformed feature space.

**Q6. What is the difference between Ridge Regression and Lasso Regression?**

* **Regularization**:
  + **Ridge Regression** uses **L2 regularization**, which penalizes the sum of the squared coefficients, but does not force coefficients to exactly zero. It shrinks coefficients towards zero without removing them.
  + **Lasso Regression** uses **L1 regularization**, which penalizes the absolute values of the coefficients and can shrink some coefficients to exactly zero, effectively performing feature selection.
* **Sparsity**:
  + Lasso tends to produce sparse models, with some features having zero coefficients.
  + Ridge tends to include all features, though with smaller coefficients.
* **Use Case**:
  + Ridge is preferred when there are many small/irrelevant features, while Lasso is preferred when we suspect that only a subset of the features are important.

**Q7. Can Lasso Regression handle multicollinearity in the input features? If yes, how?**

Yes, Lasso Regression can help address **multicollinearity** by **shrinking correlated features**. In the presence of multicollinearity, Lasso will select only one feature from a group of highly correlated variables and shrink the coefficients of others to zero. This reduces the effect of multicollinearity and produces a simpler model.

**Q8. How do you choose the optimal value of the regularization parameter (lambda) in Lasso Regression?**

The optimal value of the regularization parameter (lambda or α) in Lasso Regression is typically chosen using techniques such as:

* **Cross-validation**: Split the dataset into training and validation sets, train the model for different values of α, and evaluate the performance on the validation set. The α that minimizes validation error is selected.
* **Grid Search**: Perform an exhaustive search over a range of potential α values and evaluate the model's performance for each.
* **LassoPath or Coordinate Descent**: Algorithms like LassoPath can compute the path of coefficients as a function of α, helping to visualize the effect of different regularization strengths on the model.