

## APRENDIZAJE AUTOMÁTICO

**NOMBRE:** ALEX PÉREZ

**CÓDIGO:** 00329229

Realizar una consulta sobre las regresiones tipo Ridge, Lasso y ElasticNet donde responda al menos los siguientes tópicos:

### 1. Explique brevemente qué es Ridge Regression, Lasso Regression y ElasticNet Regression.

**Ridge Regression:** Este tipo de regresión prefiere tener un bias notable y una varianza pequeña, a tener un bias normal o pequeño y tener una varianza notable. Para esto, hablando de manera matemática, se considera un término de penalización conocido como el término L2; que representa la magnitud al cuadrado de los coeficientes. Además, también se agrega el coeficiente lambda para controlar el término de penalización, como lo que vimos en clase.

$$L_{ridge} = \operatorname{argmin}_{\hat{\beta}} \left( \|Y - \beta * X\|^2 + \lambda * \|\beta\|_2^2 \right)$$

**Lasso Regression:** Este tipo de regresión también utiliza el término de penalización, pero el significado que tiene es diferente. El término representa la suma absoluta de los coeficientes, debido a la ausencia del exponente cuadrado. De esta forma, si los coeficientes empiezan a crecer y son diferentes de 0, este término lo que hace es penalizar, ocasionando que el modelo de regresión disminuya el valor de los coeficientes para reducir las pérdidas en la función de costo.

$$L_{lasso} = \operatorname{argmin}_{\hat{\beta}} \left( \|Y - \beta * X\|^2 + \lambda * \|\beta\|_1 \right)$$

**ElasticNet Regression:** Este tipo de regresión combina el proceso de regularización tipo Lasso y tipo Ridge. Su objetivo principal es provocar un bias pequeño en aquellos casos en que el modelo de predicción muestra una gran dependencia en una variable en particular.

$$L_{elasticNet} = \operatorname{argmin}_{\hat{\beta}} \left( \hat{\beta} \right) \left( \sum \left( y - x_i^J \hat{\beta} \right)^2 \right) / 2n + \lambda \left( (1 - \alpha) / 2 * \sum_{j=1}^m \hat{\beta}_j^2 + \alpha * \sum_{j=1}^m \left\| \hat{\beta}_j \right\| \right)$$

### 2. ¿Cuál es la función de costo que minimiza cada una de las 3 formas de regresión?

El Gradient Descent es el algoritmo para minimizar la función de costo en cada una de las 3 formas de regresión.

### 3. Explique qué hiper-parámetros deben ser optimizados (*tuned*) en cada una de las 3 formas de regresión.

Aparte de los mencionados a continuación, los tres tipos de regresión lo que intentan es encontrar un balance entre bias y varianza, parámetros de la regresión que deben ser considerados para un buen modelo de Machine Learning.

**Ridge:** término de penalización, lambda, coeficientes theta; estos dentro de la función de costo.

**Lasso:** término de penalización, lambda, coeficientes theta; estos dentro de la función de costo, sin el exponente cuadrado.

**ElasticNet:** término de penalización, lambda, coeficientes theta; estos dentro de la función de costo.

**4. ¿Cómo afectan los hiper-parámetros del punto anterior al overfitting y valores de los parámetros  $\theta$  del modelo?**

Lo que el lambda y el término de penalización intenta es encontrar un balance en cuanto a dependencia de parámetros theta en el modelo. Como se vio en clase, el overfitting, causado por una gran varianza, no refleja un modelo de predicción confiable puesto que preocupa por adecuarse al dataset de prueba y no adaptarse a un modelo de predicción. De esta forma, al agregar un término de penalización, que depende de los parámetros theta, el overfitting dejaría de ser un problema, pero no se perdería la existencia de un bias o una varianza notables.

**5. ¿Por qué se dice que Lasso Regression actúa como una forma de selección de características (feature selection)?**

Porque no tiene en cuenta la característica de “multicollinearity”. Esta característica ocurre cuando algunas de las variables dependientes se encuentran correlacionadas las unas con las otras. De esta forma, Lasso ignora esta correlación existente y escoge una al azar, sin tener en cuenta su “multicollinearity”.

**6. Basado en los puntos anteriores, resume las diferencias entre las 3 formas de regresión.**

- Diferentes parámetros agregados a la función de costo como término de penalización.
- Diferente manera de balancear el bias y la varianza del modelo de regresión.
- Diferente forma de seleccionar las características del modelo, afectando a variables dependientes e independientes.

**7. Liste las clases dentro de la librería sklearn que implementan las 3 formas de regresión**

**ElasticNet:** from sklearn.ensemble import ElasticNet

**Ridge:** from sklearn import linear\_model / sklearn.linear\_model.Ridge

**Lasso:** from sklearn import linear\_model / sklearn.linear\_model.Lasso

Fuentes:

<https://www.geeksforgeeks.org/lasso-vs-ridge-vs-elastic-net-ml/>

<https://medium.com/@vijay.swamy1/lasso-versus-ridge-versus-elastic-net-1d57cfc64b58>

[https://scikit-learn.org/stable/modules/linear\\_model.html#lasso](https://scikit-learn.org/stable/modules/linear_model.html#lasso)