

L U M B R E



# Regresión regularizada

**Grado:** 5

**Materia:** Aprendizaje Supervisado · Machine Learning

11/10/2026

## Antes de empezar

Cuando hay muchas características o están correlacionadas, la regresión lineal libre se desboca: coeficientes enormes que persiguen el ruido. La regularización doma esa ansia penalizando el tamaño de los coeficientes, y de ese castigo nacen tanto la estabilidad como la selección automática de variables.

### QUÉ APRENDERÁS

- Aplicar ridge y lasso a problemas de predicción real, y comparar cómo cada penalización gestiona el sobreajuste y la selección de variables.

## Regresión regularizada

### Menos ajuste para predecir mejor

El mejor modelo no es el que ajusta mejor los datos que viste, sino el que predice mejor los que no. La regularización acepta sacrificar un poco de ajuste en entrenamiento para ganar mucho en generalización: es domar la complejidad en nombre de la honestidad.

#### VIDEO

### Regularización con Ridge y Lasso

Vídeo en español que muestra cómo penalizar el tamaño de los coeficientes frena el sobreajuste y cómo lasso llega a anular variables.



Escanea para verlo en Lumbre

<https://www.youtube.com/watch?v=uomXFxBGMvA>

## El problema de la libertad total

Una regresión lineal sin restricciones minimiza el error de entrenamiento como sea, incluso inflando coeficientes para explotar correlaciones espurias. Con muchas

variables o con columnas casi redundantes, la solución se vuelve inestable y optimista. La regularización pone un freno.

## La idea de penalizar

Se añade a la pérdida un término que crece con el tamaño de los coeficientes. Minimizar pasa a ser un equilibrio entre dos fuerzas: ajustar bien los datos y mantener los coeficientes pequeños. Un parámetro de fuerza decide cuánto peso tiene cada objetivo.

## Ridge: penalización cuadrática

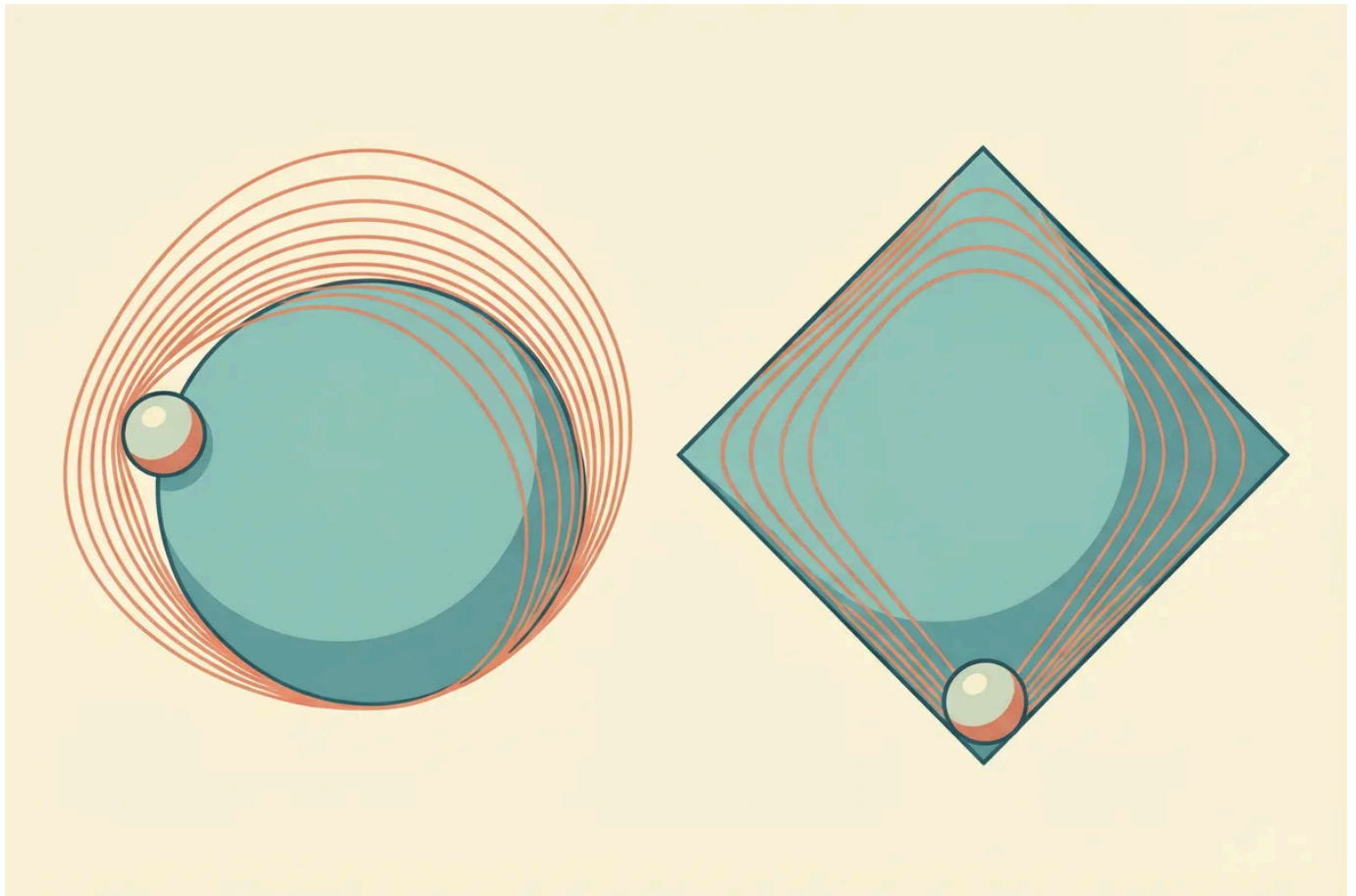
Ridge suma el cuadrado de los coeficientes —pena  $l_2$ —. Como castiga sobre todo los valores grandes, encoge todos los coeficientes hacia cero sin anularlos ninguno. Estabiliza cuando las variables están correlacionadas, repartiendo el peso entre ellas en vez de volverse loco.

## Lasso: penalización absoluta

Lasso usa el valor absoluto de los coeficientes —pena  $l_1$ —. Su geometría empuja algunos coeficientes exactamente a cero, eliminando variables del modelo. Ese escaso resultado es oro para interpretar y para elegir predictores en mares de candidatos.

## La geometría detrás

Ridge restringe la solución a un disco —suave, nunca toca los ejes—; lasso a un rombo con esquinas justo sobre los ejes, donde algún coeficiente se anula. Esa diferencia de forma explica por qué uno encoge y el otro selecciona, aunque ambos partan de la misma idea de limitar el tamaño.



*La geometría de la penalización: ridge encoge dentro de un disco sin tocar los ejes, mientras lasso, con su rombo, empuja algún coeficiente justo a cero y así selecciona variables.*

## Red elástica

Cuando hay grupos de variables correlacionadas que importan juntos, lasso elige una al azar del grupo y descarta el resto. La red elástica combina las dos penas: selecciona como lasso pero mantiene la cohesión de ridge ante predictores hermanados.

## El parámetro lambda

La fuerza de la penalización —lambda— es el hiperparámetro central. Demasiada anula señales útiles y subajusta; poca deja el descontrol original. Se elige por validación cruzada, y con ella el grado de parsimonia del modelo final.

## Escalado previo obligatorio

Penalizar el tamaño solo tiene sentido si todas las variables miden en escalas comparables: si no, la pena castiga injustamente a las de unidad pequeña. Estandarizar las columnas antes de aplicar ridge o lasso no es opcional, es un requisito para que la penalización haga lo que debe.

## Errores frecuentes

- Regularizar sin estandarizar, dejando que la escala dicte qué coeficientes sobreviven.
- Elegir lambda mirando el ajuste en entrenamiento en vez de la validación.
- Interpretar un coeficiente anulado por lasso como prueba de irrelevancia causal de la variable.

## Ejemplo resuelto: predecir consumo con cien predictores

Una distribuidora quiere estimar el consumo mensual de cada cliente a partir de casi cien variables muy correlacionadas entre sí. El procedimiento:

1. Escalar: estandarizar todas las columnas para que la penalización trate a todas por igual.
2. Elegir pena: arrancar con ridge si se sospecha que casi todo aporta; probar lasso si se cree que mandan unas pocas.
3. Sintonizar lambda: barrer valores por validación cruzada y quedarse con el que mejor predice fuera de muestra.
4. Comparar: mirar error de validación y cuántas variables sobreviven; lasso suele dejar una lista corta y legible.
5. Validar: confirmar en un conjunto apartado que el modelo no persigue el ruido.

## ¿Para qué sirve en la realidad?

La regresión regularizada es el estándar cuando hay más candidatas que muestras o están muy correlacionadas: precios inmobiliarios, riesgo clínico, demanda energética. Ridge estabiliza pronósticos; lasso entrega modelos delgados que se pueden explicar a un negocio.

### 1. ¿Qué distingue al lasso del ridge en el tratamiento de los coeficientes?

- ☐ Que el lasso nunca reduce los coeficientes
- ☐ Que el lasso puede dejar coeficientes exactamente en cero y así seleccionar variables
- ☐ Que el ridge elimina variables y el lasso no
- ☐ Que ambos son idénticos en resultado

### 2. La regularización suele empeorar el ajuste en entrenamiento para mejorar la predicción en datos nuevos.

- ☐ Verdadero
- ☐ Falso

## Cuándo usar cada uno

Con muchas variables y sospecha de que solo algunas importan, lasso y su selección vienen bien. Con predictores correlacionados que aportan en conjunto, ridge reparte el peso con más estabilidad. Ante la duda, la red elástica deja afinar la mezcla.

ACTIVIDAD INTERACTIVA DISPONIBLE EN LUMBRE

### 3. Regularización



## 4. Ordena el flujo para ajustar un modelo regularizado:

### Elementos:

- estandarizar las características a escala comparable
- elegir la penalización adecuada, l1 o l2 o mezcla
- fijar la fuerza lambda mediante búsqueda
- validar cruzada y confirmar generalización

|                |       |
|----------------|-------|
| <b>Escalar</b> | . . . |
| <b>Pena</b>    | . . . |
| <b>Lambda</b>  | . . . |
| <b>Validar</b> | . . . |

## 5. Une cada pena o herramienta con su efecto.

|    |                                                  |       |
|----|--------------------------------------------------|-------|
| 5. | encoge todos los coeficientes sin anular ninguno | ----- |
| 5. | lleva algunos coeficientes exactos a cero        | ----- |
| 5. | mezcla ambas penas para grupos correlacionados   | ----- |
| 5. | controla cuánta fuerza tiene el castigo          | ----- |

**Opciones:** ridge · lasso · red elastica · lambda

## Domar la complejidad

Regresión regularizada

Idea

penalizar tamaño

menos varianza

Ridge

pena l2

encoge sin anular

Lasso

pena l1

selecciona variables

Ajuste

estandarizar

sintonizar lambda

El mapa mental interactivo es interactivo en la versión online.



¿Qué es la regresión Ridge? <https://www.ibm.com/es-es/think/topics/ridge-regression>

Referencia en español sobre la penalización l2 y el control del sobreajuste.

**6. Con dos mil candidatas y pocas relevantes, ridge te devuelve un modelo denso e ilegible. ¿Qué cambiarías y qué esperas observar en los coeficientes?**

---



---



---



---



---

## Referencia curricular

|                 |                                            |
|-----------------|--------------------------------------------|
| <b>Grado:</b>   | 5                                          |
| <b>Materia:</b> | Aprendizaje Supervisado · Machine Learning |

Versión interactiva completa (con actividades, videos y autoevaluación):



<https://univ.lumbre.lat/object/cdia-supra-007-regresion-regularizada>

Material educativo generado desde Lumbre · [lumbre.lat](https://lumbre.lat)