

# Regresión Regularizada: Ridge y LASSO

Inferencia Estadística - UAO

## 1. Introducción

La regularización es una técnica utilizada para prevenir el sobreajuste en modelos de regresión lineal al penalizar la magnitud de los coeficientes.

- **Ridge Regression** agrega una penalización L2:  $\lambda \sum \beta_j^2$
  - **LASSO** agrega una penalización L1:  $\lambda \sum |\beta_j|$
- 

## 2. Preparación de los datos

Utilizamos el conjunto `mtcars` para predecir el consumo de combustible (`mpg`) a partir del resto de las variables.

```
data(mtcars)
X <- model.matrix(mpg ~ ., data = mtcars)[, -1]
y <- mtcars$mpg
```

---

## 3. Ajuste de modelos Ridge y LASSO

```
lambdas <- 10^seq(4, -2, length = 100)

ridge_model <- glmnet(X, y, alpha = 0, lambda = lambdas)
lasso_model <- glmnet(X, y, alpha = 1, lambda = lambdas)
```

---

## 4. Validación cruzada para elegir $\lambda$

```
set.seed(123)
cv_ridge <- cv.glmnet(X, y, alpha = 0)
cv_lasso <- cv.glmnet(X, y, alpha = 1)

cat("Mejor lambda Ridge:", cv_ridge$lambda.min, "\n")

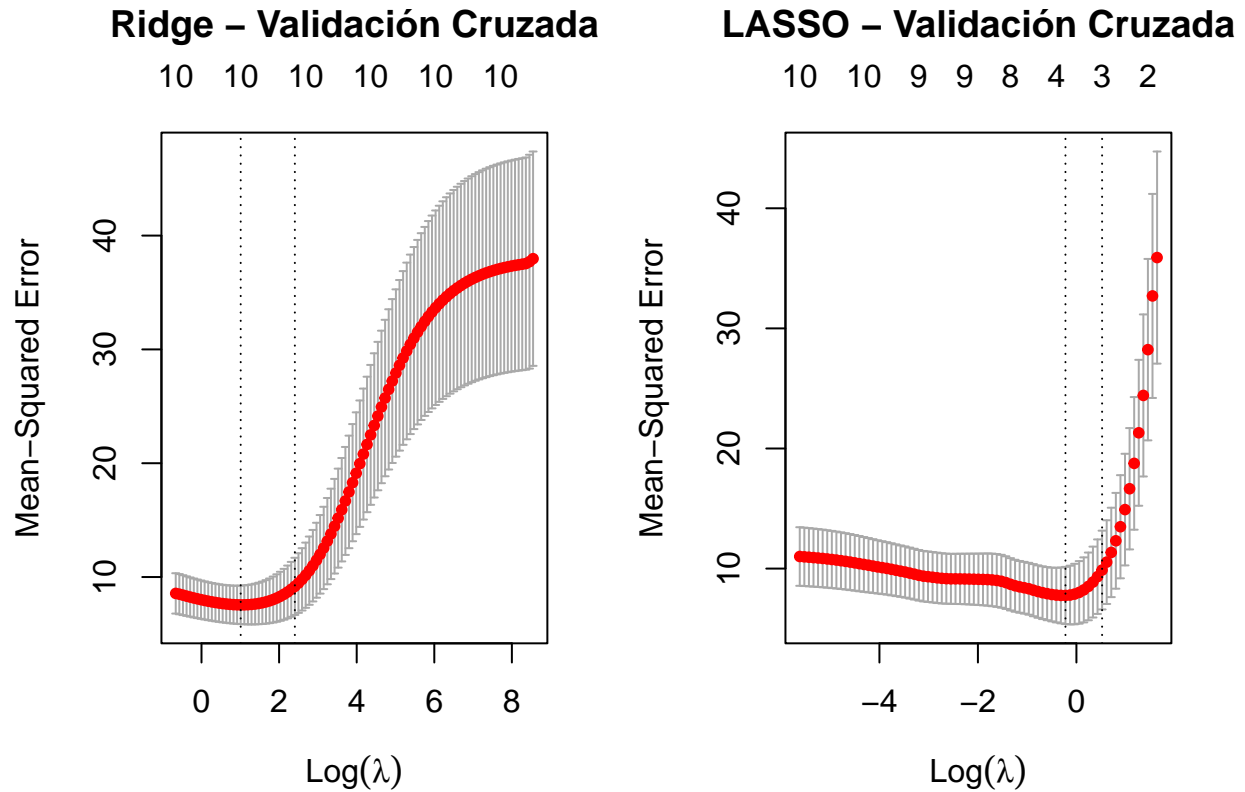
## Mejor lambda Ridge: 2.746789
cat("Mejor lambda LASSO:", cv_lasso$lambda.min, "\n")

## Mejor lambda LASSO: 0.8007036
```

---

## 5. Visualización de la validación cruzada

```
par(mfrow = c(1,2))
plot(cv_ride); title("Ridge - Validación Cruzada", line = 2.5)
plot(cv_lasso); title("LASSO - Validación Cruzada", line = 2.5)
```



## 6. Coeficientes óptimos

```
coef(cv_ride, s = "lambda.min")
```

```
## 11 x 1 sparse Matrix of class "dgCMatrix"
##              s1
## (Intercept) 21.117339411
## cyl         -0.371338786
## disp        -0.005254534
## hp          -0.011613216
## drat         1.054768651
## wt          -1.234201216
## qsec         0.162451314
## vs           0.771959823
## am           1.623812912
## gear         0.544171362
## carb        -0.547415029
```

```
coef(cv_lasso, s = "lambda.min")
```

```
## 11 x 1 sparse Matrix of class "dgCMatrix"
##              s1
```

```
## (Intercept) 36.00001676
## cyl         -0.88608541
## disp         .
## hp          -0.01168438
## drat         .
## wt          -2.70814703
## qsec         .
## vs           .
## am           .
## gear         .
## carb         .
```

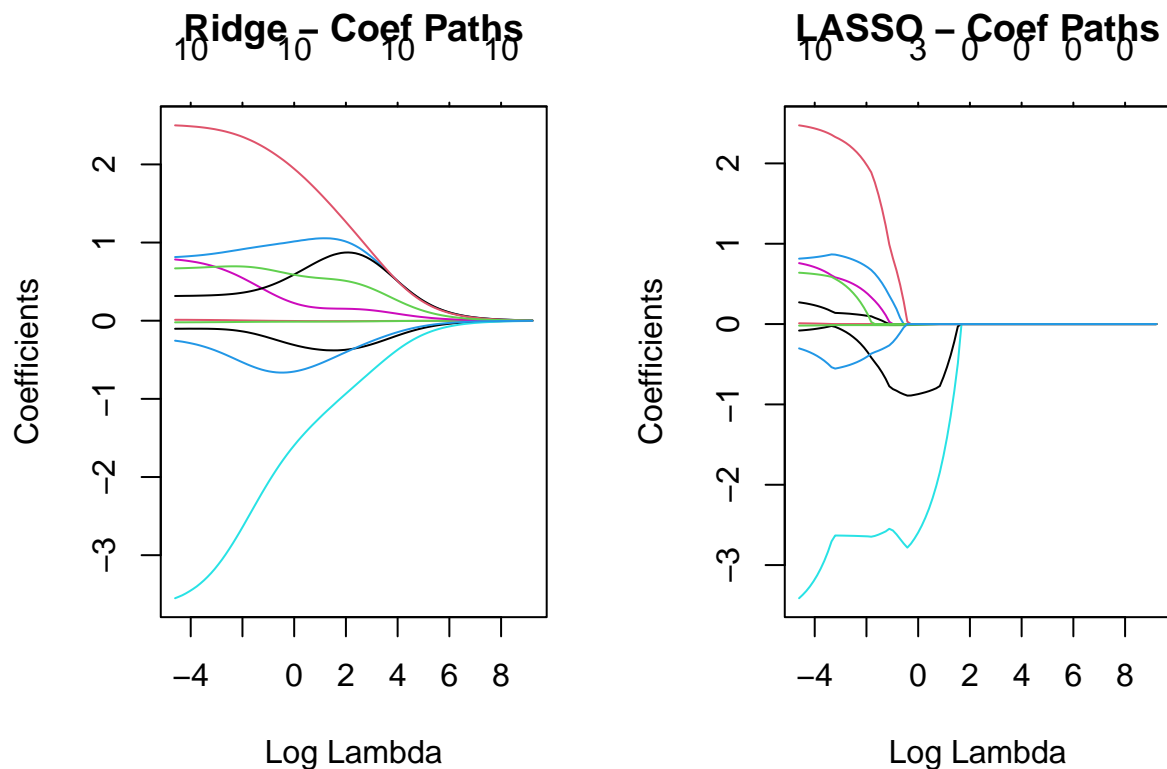
---

## 7. Interpretaciones

- Ridge reduce la magnitud de todos los coeficientes, pero **ninguno se vuelve cero**.
  - LASSO puede eliminar completamente variables (coeficientes exactamente cero), lo que permite hacer **selección automática de variables**.
  - Ambos métodos controlan la complejidad del modelo y evitan el sobreajuste.
  - La elección de  $\lambda$  es crucial y se realiza con validación cruzada.
- 

## 8. Comparación gráfica de caminos de coeficientes

```
par(mfrow = c(1,2))
plot(ridge_model, xvar = "lambda", label = TRUE)
title("Ridge - Coef Paths")
plot(lasso_model, xvar = "lambda", label = TRUE)
title("LASSO - Coef Paths")
```



## 9. Conclusión

- Usa **Ridge** cuando hay colinealidad entre variables.
- Usa **LASSO** si quieres simplificar el modelo y seleccionar variables.
- Ambos son fundamentales en la estadística moderna para análisis en alta dimensión.

## 10. Ejemplo en Alta Dimensión con Datos Reales: Leucemia

Usaremos datos del paquete `glmnet`, que contiene mediciones de expresión génica en pacientes con leucemia (AML vs ALL). Hay: -  $n = 72$  observaciones (pacientes) -  $p = 7129$  predictores (genes)

```
load("/Users/cristiangarcia/Dropbox/Docencia/UA0/2025_1-UA0/Inferencia_maestria/Clases/Sesion_8/golub.RData")
X <- t(golub) # X debe ser n x p (72 x 3051)
y <- factor(golub.cl) # 0 = ALL, 1 = AML

# Ajuste de modelos
ridge_leukemia <- cv.glmnet(X, y, alpha = 0, family = "binomial")
lasso_leukemia <- cv.glmnet(X, y, alpha = 1, family = "binomial")
```

## 11. Resultados de clasificación

```
cat("Error CV Ridge:", min(ridge_leukemia$cvm), "\n")
```

```
## Error CV Ridge: 0.3518916
```

```
cat("Error CV LASSO:", min(lasso_leukemia$cvm), "\n")
```

```
## Error CV LASSO: 0.1501018
```

---

## 12. Número de genes seleccionados por LASSO

```
coef_lasso_leukemia <- coef(lasso_leukemia, s = "lambda.min")
selected_genes <- sum(coef_lasso_leukemia[-1, ] != 0)
cat("Genes seleccionados por LASSO:", selected_genes)
```

```
## Genes seleccionados por LASSO: 14
```

---

## 13. Conclusión del Caso de Alta Dimensión

- Ridge logra un buen ajuste, pero no realiza selección.
- LASSO logra error comparable o menor, **eliminando miles de genes irrelevantes**.
- Ideal para problemas tipo ‘Big P, Small n’, como genómica o finanzas.

Este ejemplo muestra el poder de los métodos de regularización para trabajar en escenarios donde el número de predictores supera ampliamente al número de observaciones.