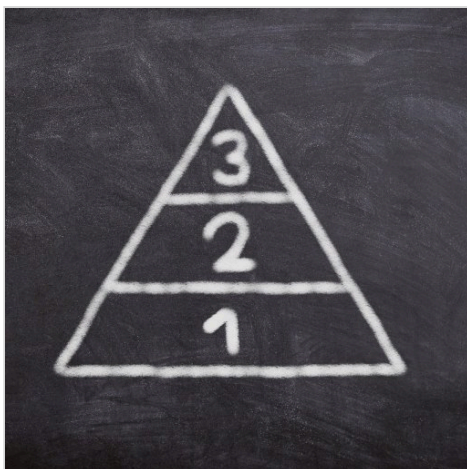


L U M B R E



Bondad de ajuste y selección de modelos

Grado: 4

Materia: Regresión y Modelado Estadístico · Probabilidad y Estadística

11/10/2026

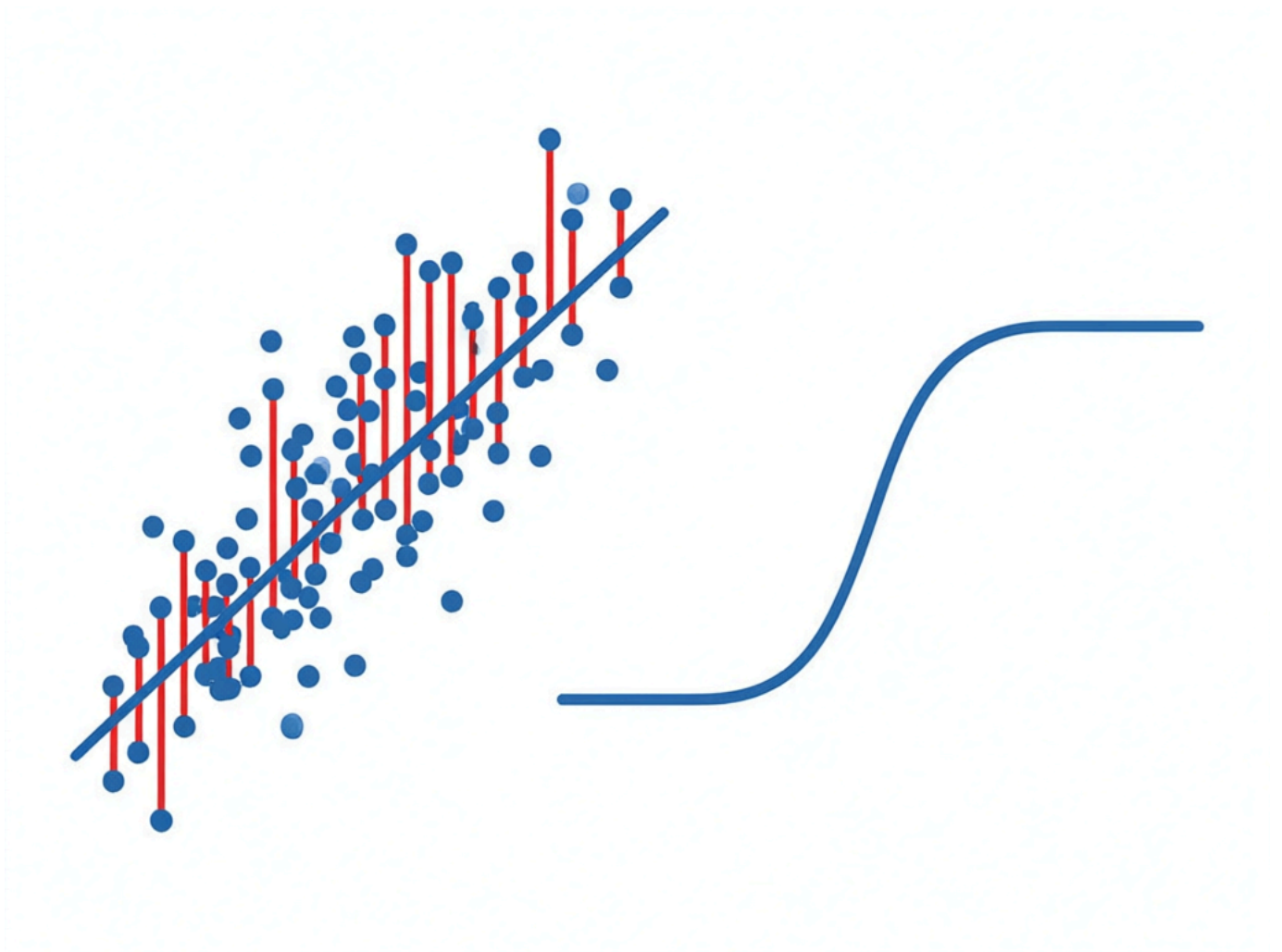
Antes de empezar

Cuando hay muchos modelos plausibles, ¿cuál elegir? La tentación de quedarse con el que más acierta en los datos disponibles lleva al sobreajuste. Esta lección enseña a comparar modelos penalizando la complejidad y validando de verdad.

QUÉ APRENDERÁS

- Evaluar y comparar modelos con criterios de verosimilitud, información y validación, y aplicar métodos de selección de variables sin engañarse con el ajuste en el entrenamiento.

Bondad de ajuste y selección de modelos



Regresión y modelado estadístico: lineal, múltiple y logístico.

Mapa conceptual

- **Regresión**

- Lineal simple
 - Beta 0 y 1
 - R-cuadrado
 - Residuos
- Supuestos
 - Normalidad
 - Homocedasticidad
 - Independencia
- Múltiple
 - Multicolinealidad
 - Ajuste R^2
 - AIC/BIC
- Logística
 - Odds ratio
 - Sigmoide
 - Regularización

El mapa mental interactivo es interactivo en la versión online.



No ganes el concurso equivocado

El modelo que mejor describe los datos con que te ajustaste casi nunca es el que mejor verá los datos del mañana. Seleccionar modelos es el arte de preferir uno que generalice, no uno que adule.

El problema de comparar modelos

Añadir predictores nunca empeora el ajuste en el entrenamiento: la bondad medida ahí sube monótona. Si eligiéramos por ese criterio, siempre ganarían los modelos más complejos, que suelen estar más sobreajustados. Comparar exige un correctivo.

Verosimilitud y su trampa

La verosimilitud mide cuán bien un modelo explica los datos observados, y también sube al añadir parámetros. Necesitamos criterios que descuezan la ganancia de bondad contra el coste de la complejidad añadida, para no premiar el simple hecho de tener más grados de libertad.

Criterios de información

El AIC y el BIC resumen verosimilitud menos una penalización por número de parámetros. Premian el buen ajuste pero cobran peaje por cada coeficiente. El mejor modelo es el que minimiza el criterio, equilibrando acierto y parsimonia.

AIC frente a BIC

El AIC penaliza menos los parámetros y tiende a modelos más ricos, orientado a predicción. El BIC penaliza más, sobre todo con muestras grandes, y favorece modelos más simples, apuntando a recuperar la estructura verdadera. Eligen según el objetivo y el tamaño.

Validación como juez último

Ningún criterio interno sustituye a medir el error sobre datos no usados en el ajuste. La validación cruzada estima cómo rendirá el modelo ante datos nuevos, que es justo lo que queremos, y es el árbitro más honesto entre candidatos.

Validación cruzada k pliegues

Se parte la muestra en k pliegues; cada vez se entrena con k menos uno y se valida con el restante. El error promedio estima la capacidad de generalizar. Repetir con distintas particiones reduce la variabilidad de esa estimación.

Selección por pasos

La selección hacia delante añade la variable que más mejora en cada paso; la hacia atrás parte de todas y elimina. Ambas son codiciosas y frágiles: pueden quedarse enganchadas en elecciones tempranas mediocres, y se desaconsejan frente a la regularización moderna.

Mejor subconjunto

Evaluar todas las combinaciones de predictores daría el óptimo pero es exponencialmente inviable. Los métodos voraces la imitan de forma barata, a costa de poder perder el mejor modelo. La regularización ofrece hoy una alternativa más estable.

R square y su prima ajustada

En modelos anidados, el incremento de R cuadrada al añadir una variable puede deberse al azar. La R cuadrada ajustada y los tests de comparación de modelos

anidados ayudan a juzgar si el extra merece la pena, penalizando la ganancia superficial.

El peligro del múltiple comparado

Probar decenas de modelos y quedarse con el mejor según el mismo conjunto de validación reintroduce el sobreajuste: se selecciona el que tuvo más suerte. Un conjunto de test reservado hasta el final es la única salvaguardia contra la optimización involuntaria.

Criterios para clasificación

Cuando la respuesta es categórica, la bondad no se mide con residuos cuadrados sino con verosimilitud, entropía o curvas ROC. Los criterios de información se construyen entonces sobre la verosimilitud, manteniendo la misma lógica de penalizar la complejidad.

Curva de aprendizaje

Dibujar el error de entrenamiento y validación contra el tamaño de muestra diagnostica el tipo de problema: brecha amplia entre ambas señala sobreajuste (más datos ayudarían); ambas altas y juntas, subajuste (hace falta un modelo más expresivo).

Errores frecuentes

- Elegir el modelo por su ajuste en el entrenamiento, sin validación.
- Probar muchos modelos y quedarse con el mejor según el mismo conjunto de validación.
- Premiar siempre la complejidad ignorando la parsimonia y la interpretabilidad.

1. El motivo principal por el que el AIC penaliza el número de parámetros es:

- ☐ Hacer el modelo más lento
- ☐ Evitar premiar la complejidad que solo mejora el ajuste en el entrenamiento
- ☐ Garantizar la linealidad
- ☐ Aumentar la verosimilitud

2. La validación cruzada estima directamente el error de un modelo sobre datos no vistos.

- ☐ Verdadero
- ☐ Falso

Parsimonia y sentido común

Un modelo ligeramente peor en métricas pero que un experto entiende y confía, a menudo vale más en producción que una caja negra opaca con una décima más de acierto. La selección no es solo numérica: la interpretabilidad tiene valor de uso.

El modelo es un medio

Seleccionar modelo no es un concurso de métricas sino un medio para una decisión. Definir primero qué error cuesta más, y elegir criterio y validación coherentes con ese coste, importa más que rascar una milésima de acierto con el criterio equivocado.

ACTIVIDAD INTERACTIVA DISPONIBLE EN
LUMBRE

3. Selección de modelos



4. Ordena un proceso honesto de selección:

Elementos:

- proponer modelos plausibles
- evaluar cada uno con validación cruzada
- comparar con criterios que penalicen la complejidad
- confirmar en un test jamás mirado

Candidatos	. . .
Validar	. . .
Comparar	. . .
Reservar	. . .

Interpretable ML <https://christophm.github.io/interpretable-ml-book/Comparación y selección de modelos con criterio>.

VIDEO

Video complementario

Recurso audiovisual para reforzar los conceptos.



Escanea para verlo en Lumbre

<https://www.youtube.com/watch?v=1wK1K2M7tXI>

5. Ajustas doce modelos y el mejor gana por poco en validación cruzada. Antes de elegirlo, ¿qué comprobarías para no estar sobreoptimizando la selección?

NIST: Regression <https://www.itl.nist.gov/div898/handbook/pmd/section1/pmd16.htm>
Referencia completa de metodos de regresion.

Referencia curricular

Grado:	4
Materia:	Regresión y Modelado Estadístico · Probabilidad y Estadística

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



<https://univ.lumbre.lat/object/cdia-reg-006-bondad-de-ajuste-y-seleccion-de-modelos>

Material educativo generado desde Lumbre · lumbre.lat