Regresión polinómica (no lineal) (polynomial regression) (Página 290)

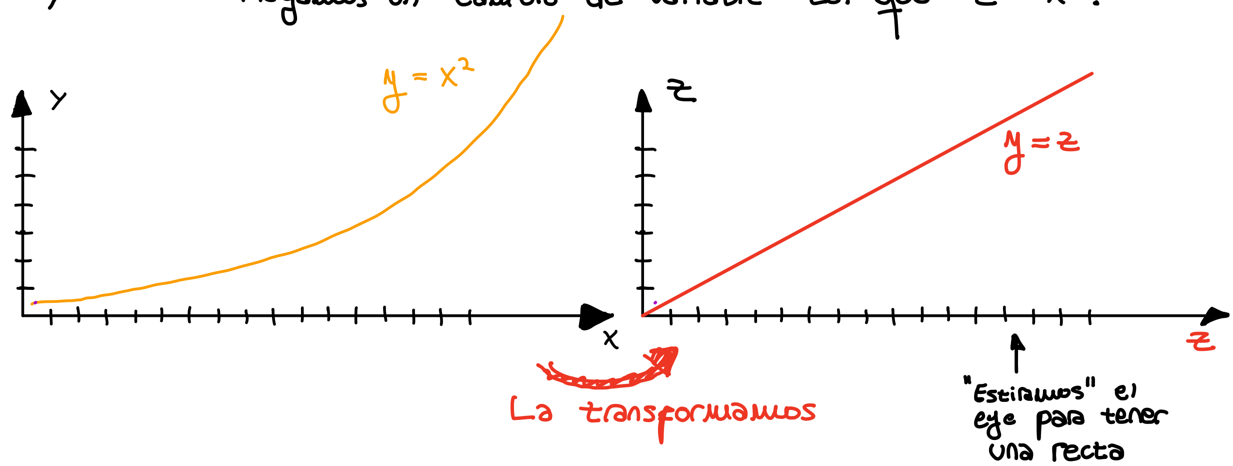
No siempre tendremos relaciones lineales. Tendremos que ir a las polinómicas. Pasamos a tener el siguiente modelo (veámoslo con una sola variable).

$$\boxed{y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_d x_i^d + \varepsilon_i}$$

Polinomio intercalador de Lagrange

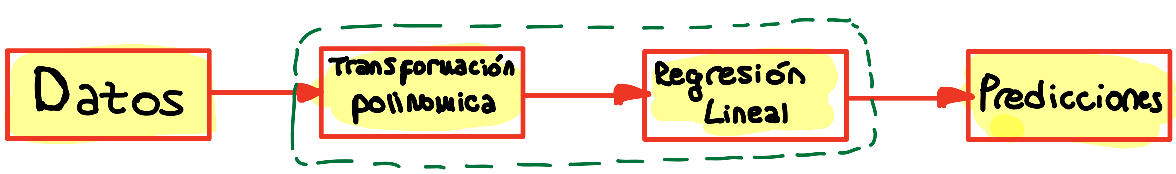
"grado" → Se usa hasta 4. Pasando eso es muy raro que se use

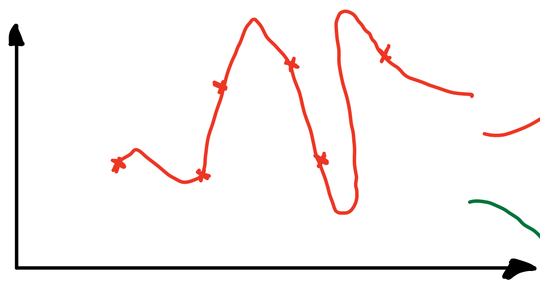
¿Cómo trabajamos con esto? Supongamos que tenemos la función $y = x^2$. Hagamos un cambio de variable tal que $z = x^2$:



Transformamos lo lineal en lineal. ¿Para qué? Para trabajar con regresión lineal.

En vez de trabajar con $y = \beta_0 + \beta_1 x^2$, trabajamos con $y = \beta_0 + \beta_1 z$. Realizamos una transformación a cada variable. El R^2 lo calculamos entre y y z, no entre x e y.





Si el ajuste pasa por todos los puntos $(d = p-1)$ NO HAY ERROR, lo cual esta mal! No hay aprendizaje.

Sobre ajuste

¿Como evitamos el sobre ajuste? Tendriamos que sacarle dependencia a la variable x_i ($\beta_i \to 0$). Evitamos sobreajuste.

A esto se lo llama Regularización.

(shrinkage, regularization)

se le agrega un peso

Regularización → forma de regresión en el que lo coeficientes tenderan a ir a 0.

Nosotros queriamos minimizar SCE que es

$$SCE = \sum_{i=1}^{n} \left[y_i \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right]^2$$

Lo que hacemos ahora es minimizar

Tuning parameter

Ridge Regression

Cuanto "penalizamos" a la flexibilidad del modelo

$$SCE + \lambda \sum_{j=1}^{p} \beta_j^2$$

"Mas te vale que la suma sea chica"

Le agregamos exigencia

Realmente puede hacer que β_j sea 0.

El metodo del Lasso quiere que se tome $|\beta_j|$ y no β_j^2

Consideraciones de λ

- Si $\lambda \to \infty$, todos los β_j sean 0
- Si $\lambda \to 0$, es como si no hubiera penalización

- R^2 ajustado → Penaliza una métrica
- Regularización → Penaliza un modelo

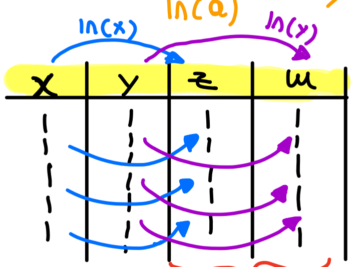
Si queremos estimar sobre cualquier función, siempre linealizamos. Los pasos a seguir son:

1) Linealizamos la función propuesta
2) Transformamos los datos
3) Hacemos regresión lineal
4) Hacemos los despeyes para encontrar los parámetros de la función

Eyemplos

$$y = a \cdot x^b$$

1) $\ln(y) = \ln(a) + b \cdot \ln(x)$
2) $w = \underbrace{\beta_0}_{\ln(a)} + \underbrace{\beta_1}_{\beta} \cdot \underbrace{z}_{\ln(x)}$ El modelo queda aditivo



Se desarrollan estas tablas

3) Hacemos regresión lineal (β_0 y β_1)
4) Despeyamos a y b

- $a = e^{\beta_0}$
- $b = \beta_1$