Métodos de construcción de los coeficientes

Regularización

Se tiene un modelo de la forma . Se define el error residual como y se busca .

La función es la función de costo .

La regularización consiste en modificar sumándole un , donde es un factor que depende de la magnitud de (o sea de su norma). La mera función de costo es . es la constante de regularización.

Ejemplos:

1. Lasso
2.  Ridge: Aquí hay un abuso de notación, porque el no se regulariza.

Cuando no hay regularización.

Cuando hay regularización.

penaliza las variables con coeficientes grandes. Si la variable no reduce el error entonces se castiga que tenga un coeficiente grande.



Supongamos que



Con predictores anlados se tiene

Nota: Tanto ||.||1 como ||.||2 son sensibles a la escala. Esta aproximación solo debe usarse con variables estandarizadas (Centradas y escaladas).

¿Por qué Lasso es el mejor para selección de variables? Porque escoge rápido los coeficientes.



Denotemos por los estimadores Ridge y Lasso para un valor particular de y el estimador para (Mínimos cuadrados o máxima verosimilitud).

es el estimador para el i-ésimo predictor (que debe estar estandarizada).

