

Regularizacion B. $p > n$. Elastic Net

Test de Lectura

Como discutiesemos, estos *tests de lectura* son una serie de preguntas simples que permiten monitorear si estan siguiendo efectivamente las clases. No seran corregidos ni evaluados, los administraran directamente ustedes. Cualquier duda que tengan sera discutida en nuestras sesiones sincronicas, la semana siguiente a publicado el material, luego de que hayan hecho un esfuerzo en trabajarlo, no antes. Es muy importante que trabajen este material y lo discutan entre ustedes.

1. Repasa las condiciones de KKT para el problema de minimizacion restringida de ridge.
2. Chequea que si tomamos $C = \hat{\beta}(\lambda)^2$, $\alpha = \lambda$ y $\hat{\beta}(\lambda)$ satisfacen la condiciones de KKT.
3. Mira bien el dibujo del slide 34 y lee la discusion del libro, de la pagina 222.
4. Como seria el conjunto de restriccion (para LASSO y ridge) si $\lambda = 0$. Y si $\lambda = \infty$?
5. Por que C y λ guardan una relacion negativa?
6. V o F: ridge es minimos cuadrados ordinarios sujeto a la restriccion de que los coeficientes tienen que ser mayores a cierto valor.
7. Dificil: mirando bien el dibujo del slide 34. Por que ridge no elimina coeficientes y LASSO si?
8. Por que decimos que LASSO resuelve tanto un problema de 'model selection' como de mejorar el error cuadratico medio predictivo?
9. Por que ridge resuelve el segundo problema pero no el primero?
10. Revisa bien que es el rango de una matriz. Explica por que MCO no funciona cuando $p > n$.
11. Mirando el argumento de 'data augmentation' explica por que ridge sigue funcionando aun cuando $p > n$.
12. Dificil (en realidad, molesto): para el caso de un solo predictor estandarizado. A mano alzada, 1) dibuja n puntos 2) dibuja la recta de MCO 2) Agrega el punto adicional que sugiere el razonamiento de 'data augmentation'. 3) Fijate que pasa con la recta de regresion a medida que

λ aumenta. Usa este dibujo para explicar en que sentido ridge ‘achica’ coeficientes.

13. Que problema viene a resolver elastic net?
14. Por que elastic net elige coeficientes y ridge no?
15. Obtene la solucion explicita de naive elastic net del caso del slide 8.
16. Obtene la solucion explicita pero para el caso en que $\hat{\beta}_{mco} < 0$
17. V o F: cuando $\lambda_2 = 0$, naive elastic net coincide con LASSO.