## Temperatura Crítica de Superconductores

Presentamos nuevos modelos

Grupo A - Estadística

Pontificia Universidad Católica de Chile Facultad de Matemáticas EYP2307 - Análisis de Regresión

1 de Diciembre de 2020





## Contenido

- Avance 1
- Nuevos modelos
- Elegimos modelo
- Ridge Regression
- Conclusiones
- Referencias bibliográficas





## Contenido

#### Avance 1





#### Recursos Utilizados

- Usamos RStudio.
- 2. R Markdown y R Sweave.
- GitHub.
- 4. Bases de datos.
  - train.csv
  - unique\_m.csv





## Objetivo Avance 1

 Predecir la temperatura crítica de los superconductores, en base a nuestra variable respuesta critical\_temp.





## Limpieza de la base de datos

► Como se tenían **169** variables en total, se decidió limpiar la base de datos.

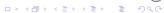




## Limpieza de la base de datos

- Como se tenían 169 variables en total, se decidió limpiar la base de datos.
- ► Al hacer la limpieza nos quedamos solo con **34** variables.





Se hizo un análisis de correlación.





- Se hizo un análisis de correlación.
- La variable std\_ThermalConductivity tuvo la correlación más alta de 0.65, por lo tanto se utilizó para nuestro modelo.





- Se hizo un análisis de correlación.
- La variable std\_ThermalConductivity tuvo la correlación más alta de 0.65, por lo tanto se utilizó para nuestro modelo.
- Al hacer el análisis de la varianza explicada:  $R^2 = 0.43$ .





- Se hizo un análisis de correlación.
- La variable std\_ThermalConductivity tuvo la correlación más alta de 0.65, por lo tanto se utilizó para nuestro modelo.
- ▶ Al hacer el análisis de la varianza explicada:  $R^2 = 0.43$ .
- Se decidió buscar alternativas para intentar aumentar este último valor.





Nos decidimos por un nuevo modelo.





- Nos decidimos por un nuevo modelo.
- Utilizamos la variable range\_Valence por ser una variable discreta y nos quedaron 6 modelos.





- Nos decidimos por un nuevo modelo.
- Utilizamos la variable range\_Valence por ser una variable discreta y nos quedaron 6 modelos.
- El modelo final nos queda:





- Nos decidimos por un nuevo modelo.
- Utilizamos la variable range\_Valence por ser una variable discreta y nos quedaron 6 modelos.
- El modelo final nos queda:
  - ightharpoonup Correlación = **0.75**.





- Nos decidimos por un nuevo modelo.
- Utilizamos la variable range\_Valence por ser una variable discreta y nos quedaron 6 modelos.
- ► El modelo final nos queda:
  - Correlación = 0.75.
  - $R^2 = 0.56.$





## Objetivo del Avance 2

Predecir la temperatura crítica de los superconductores en base a nuestra variable respuesta, utilizando modelos de regresión lineal múltiple para mejorar los resultados obtenidos en el Avance 1.





## Contenido

Nuevos modelos





- Creamos una serie de nuevos modelos:
  - 1. Backward.





- Creamos una serie de nuevos modelos:
  - 1. Backward.
  - 2. Forward.





- Creamos una serie de nuevos modelos:
  - 1. Backward.
  - 2. Forward.
  - 3. Backward-Forward.





- Creamos una serie de nuevos modelos:
  - 1. Backward.
  - 2. Forward.
  - 3. Backward-Forward.
  - 4. add1.





- Creamos una serie de nuevos modelos:
  - 1. Backward.
  - 2. Forward.
  - 3. Backward-Forward.
  - 4. add1.
  - drop1.





- Creamos una serie de nuevos modelos:
  - 1. Backward.
  - 2. Forward.
  - 3. Backward-Forward.
  - 4. add1.
  - 5. drop1.
  - 6. VIF.





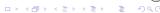
- Creamos una serie de nuevos modelos:
  - 1. Backward.
  - 2. Forward.
  - 3. Backward-Forward.
  - 4. add1.
  - 5. drop1.
  - 6. *VIF*.
  - 7. Regresión Multiple con la idea del Avance 1.





- Creamos una serie de nuevos modelos:
  - 1. Backward.
  - 2. Forward.
  - 3. Backward-Forward.
  - 4. add1.
  - 5. drop1.
  - 6. VIF.
  - 7. Regresión Multiple con la idea del Avance 1.
  - 8. Ridge Regression.





- Creamos una serie de nuevos modelos:
  - 1. Backward.
  - 2. Forward.
  - 3. Backward-Forward.
  - 4. add1.
  - 5. drop1.
  - 6. VIF.
  - 7. Regresión Multiple con la idea del Avance 1.
  - 8. Ridge Regression.
  - 9. Lasso Regression.





## Backward

▶ abc





## **Forward**

▶ abc



## Backward-Forward

Nuevos modelos 00000000000

abc





add1

▶ abc



▶ abc



VIF

► abc



## Regresión Multiple con la idea del Avance 1

abc





## Ridge Regression

Nuevos modelos

▶ abc





abc



### Contenido

Avance :

Nuevos modelo

Elegimos modelo

Ridge Regression

Conclusiones

Referencias bibliográficas





00

## Diapositiva

abc





Ridge Regression





00

# Diapositiva

abc





### Contenido

Conclusiones





abc





## Contenido

Referencias bibliográficas





## Referencias bibliográficas

- archive.ics.uci.edu/ml/datasets/Superconductivty+Data Kam Ham Idieh - Machine Learning Repository. 2018
- https://arxiv.org/pdf/1803.10260.pdf Joe Ganser - Superconductivity Regression. 2019
- Machine learning modeling of superconducting V. Stanev, C. Oses, A.G. Kusne, et al. 2018



