

LABORATORIO DE DATOS

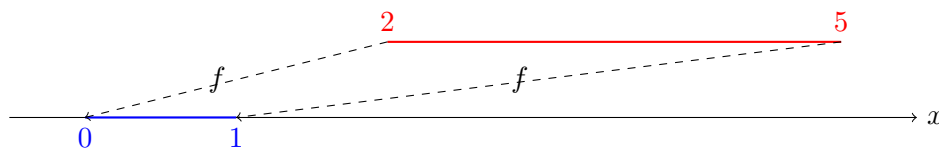
Primer Cuatrimestre 2025

Práctica N° 6: Modelo lineal multivariado y transformaciones de datos.

1. **Variables de oferta.** Utilizamos el dataset `macro_full_columns.csv`.

- (a) Generar un DataFrame que contenga solo las variables que terminan con `.oferta`. ¿Había datos faltantes en estas columnas?
- (b) Queremos explicar la variable `PBI_a_precios_de_mercado.oferta` utilizando el resto de las variables de oferta. Crear un DataFrame **X** que contenga todas las variables de oferta excepto la de PBI y una Series **y** que contenga solo la variable de PBI.
- (c) Ajustar un modelo lineal y generar el vector de predicciones.
- (d) Graficar, en un mismo gráfico, la variable respuesta original y la predicha en función del año. Sugerencia: prestar atención a los índices de cada variable.

2. Hallar un escalamiento lineal $f(x)$ que transforme los valores en el intervalo $[2, 5]$ a valores en el intervalo $[0, 1]$.



3. Para este ejercicio utilizar los datos `CollegeBasketballPlayers2009-2021.csv`.

- (a) Ajustar un modelo de regresión para predecir los puntos promedio (`pts`) en función de las demás variables. Comparar los coeficientes del modelo obtenido sin escalar variables con los coeficientes que se obtienen si primero se escalan las variables utilizando escalamiento MinMax.
- (b) A partir de los resultados, según los coeficientes obtenidos, ¿cuáles son las 5 variables de mayor peso en el modelo?