Section 3J. Interaction Terms

Statistics for Data Science

Victor M. Preciado, PhD MIT EECS
Dept of Electrical & Systems Engineering
University of Pennsylvania
preciado@seas.upenn.edu

Nonlinear effects: Interaction terms

- Consider another dataset where the output variable is the number of sales of a product in 200 different markets. The input variables are the amount invested in TV and Radio advertising in each market; hence, we can build a linear model

$$\text{Sales} = \beta_0 + \beta_1 \times \text{TV} + \beta_2 \times \text{Radio} + \varepsilon$$

  - This model assumes that the effects of TV and Radio investments are independent of each other
  - However, spending money on Radio advertising may increase the effectiveness of TV advertising
  - In statistics, the synergies between variables are called *interaction* effects and can be modeled by introducing new artificial variables in the model, as follows

$$\text{Sales} = \beta_0 + \beta_1 \times \text{TV} + \beta_2 \times \text{Radio} + \beta_3 \times (\text{TV} \times \text{Radio}) + \varepsilon$$

Nonlinear effects: Interaction terms (cont.)

- After performing a linear fitting with the output variable $y_i = \text{Sales}_i$ and the input vectors $\mathbf{x}_i = [\text{TV}_i, \text{Radio}_i, \text{TV}\times\text{Radio}_i]$, we obtain the following table of coefficients:

[Credit: James et al, ISL book]	Coefficient	Std. Error	t-statistic
Intercept	6.7502	0.248	27.23
TV	0.0191	0.002	12.70
radio	0.0289	0.009	3.24
TV×radio	0.0011	0.000	20.73

- **Interpretation**: The interaction between Radio and TV advertising is important (t-statistic is large).

Interaction terms: The Hierarchy Principle

Hierarchy Principle: If we include an interaction term in our model (e.g., TV$\times$Radio), we should also include the main terms (e.g., TV and Radio), even if the t statistics indicate that those main terms are not significant.

- The reason for this principle is that the interaction terms are hard to interpret if we do not include the main terms
- In particular, the interaction term could carry the effect of main terms if these are not included in the model

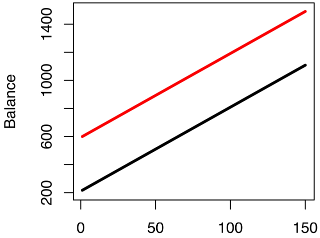
Interactions between Qualitative and Quantitative Variables

Consider an example in we want to predict the **Balance** of an individual using as inputs his/her **Income** (quantitative) and if he/she is a **Student** (qualitative)

- Model without interactions:

$$\widehat{\text{Balance}}_i = \beta_0 + \beta_1 \times \text{Income}_i + \beta_2 \times \text{Student}_i = \begin{cases} (\beta_0 + \beta_2) + \beta_1 \times \text{Income}_i, & \text{if } i \text{ is a student} \\ \beta_0 + \beta_1 \times \text{Income}_i, & \text{if } i \text{ is not a student} \end{cases}$$

- Model plots: With no interactions, we have a change in the intercept



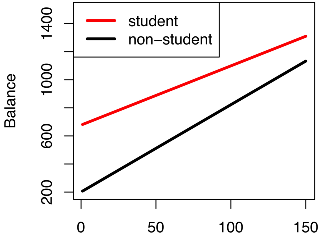
Interactions between Qualitative and Quantitative Variables (cont.)

- Model *with* interactions:

$$\widehat{\text{Balance}}_i = \beta_0 + \beta_1 \times \text{Income}_i + \beta_2 \times \text{Student}_i + \beta_3 \times \text{Income}_i \times \text{Student}_i$$

$$= \begin{cases} (\beta_0 + \beta_2) + (\beta_1 + \beta_3) \times \text{Income}_i, & \text{if } i \text{ is a student} \\ \beta_0 + \beta_1 \times \text{Income}_i, & \text{if } i \text{ is not a student} \end{cases}$$

- Model plots: With no interactions, we have a change in the intercept



Penn Engineering