Interaction Models
(Continuous x Continuous)

LI(✘) – Relationship between Y and X's are not linear

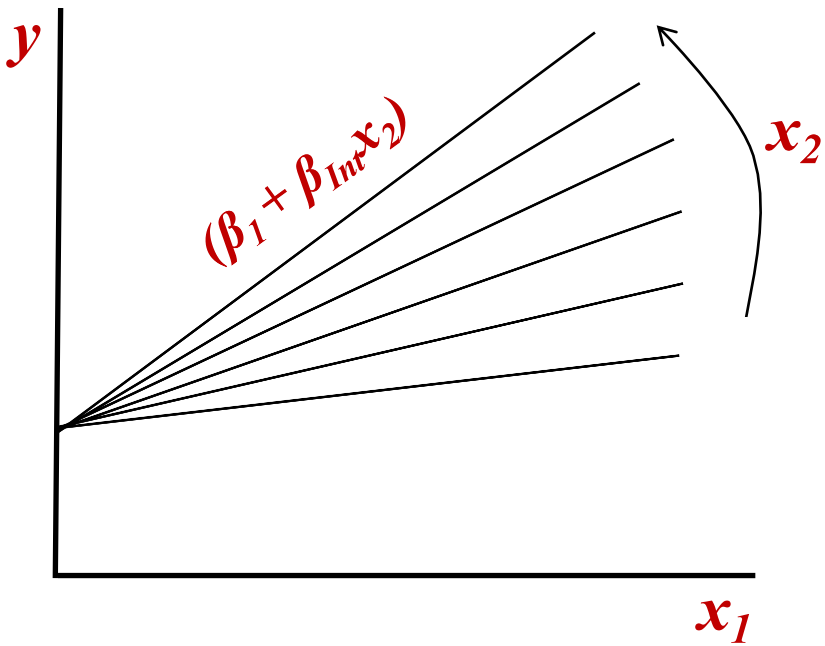
Interactions with Continuous Variables

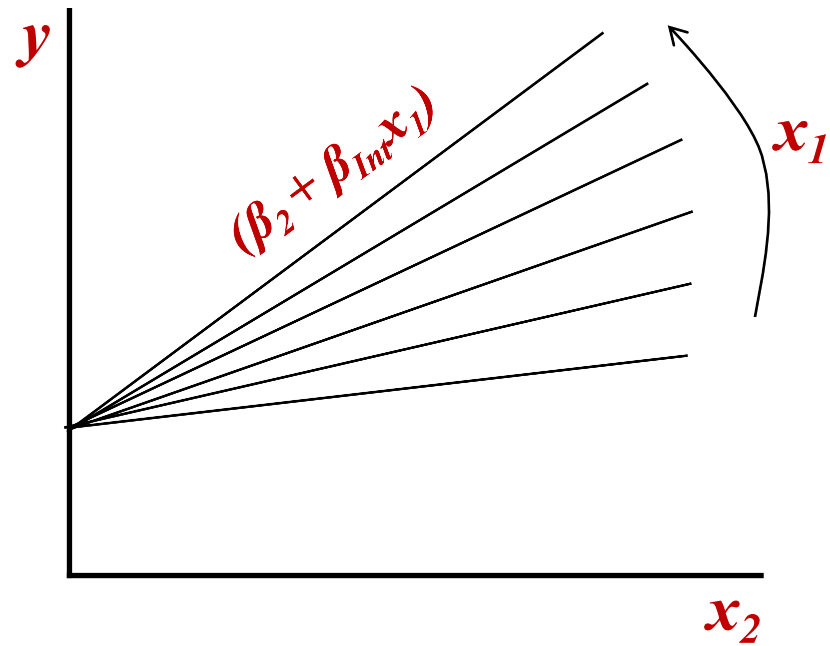
- Interactions with 2 **continuous** variables are modeled in a similar fashion than with 1 continuous x 1 dummy variable
- But they are more **challenging** (e.g., high multicollinearity) and **harder** to **interpret**, therefore less useful in predictive modeling
- Take for example: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{Int} x_1 x_2 + \varepsilon$
- Where x_1 and x_2 are both **continuous**, therefore:
  - β_1 is the effect of x_1 when x_2 is 0
  - β_2 is the effect of x_2 when x_1 is 0
  - The total effect of x_1 is $\beta_1 x_1 + \beta_{Int} x_1 x_2 = (\beta_1 + \beta_{Int} x_2)\, x_1$
  - That is, the effect of x_1 changes for every value of x_2
  - Similarly, the total effect of x_2 is $\beta_2 x_2 + \beta_{Int} x_1 x_2 = (\beta_2 + \beta_{Int} x_1)\, x_2$
  - That is, the effect of x_2 changes for every value of x_1, so

Graphical Illustration

When x_1 and x_2 are continuous

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{Int} x_1 x_2 + \varepsilon$$





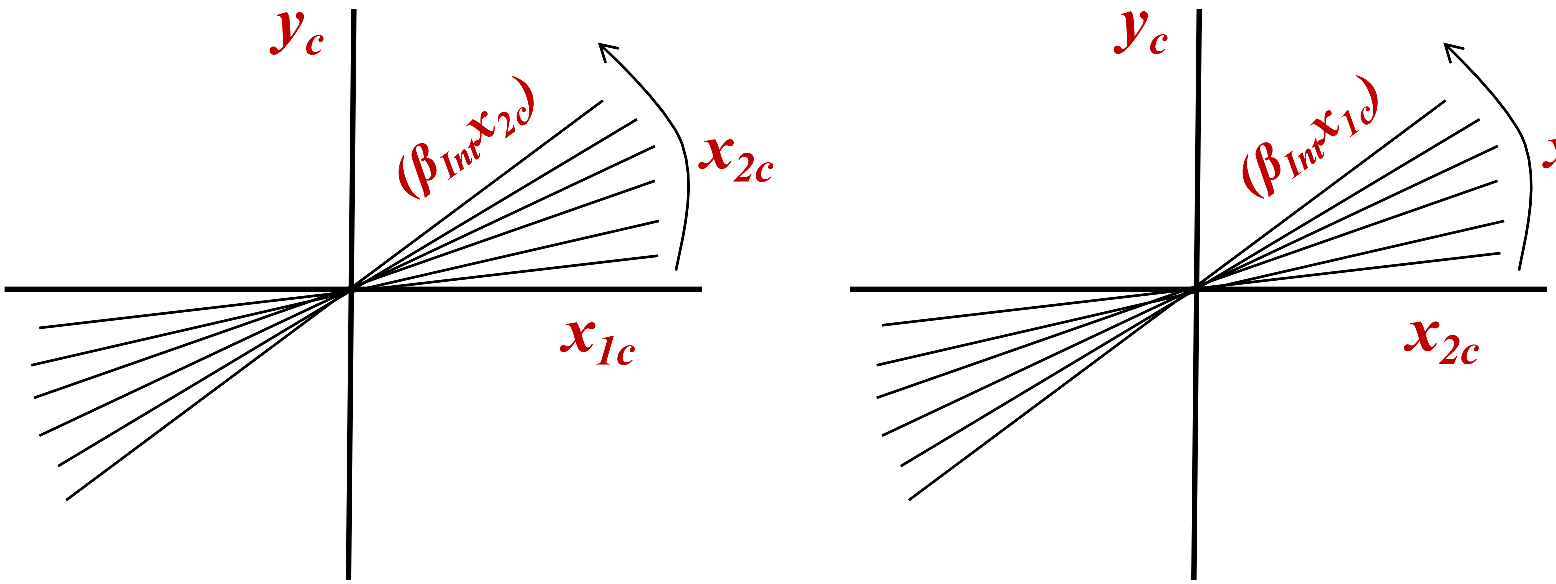
Problems with Continuous Interactions

- The model is **no** longer **scale invariant**. Since the effect of x_1 depends on x_2, re-scaling x_2 will affect the effect of x_1, and vice versa
- Multicollinearity increases substantially – i.e., if x_1 is correlated with x_2, it will also be correlated with $x_1 x_2$
- Since β_1 is the effect of x_1 when x_2 is 0, it will be meaningless if $x_2 = 0$ can never happen (e.g., vehicle weight = 0)
- **Solution: center** or **standardize** the **dependent** variable and all **independent** variables associated involved in the interaction
  - ✓ It **reduces multi-collinearity** substantially
  - ✓ β_1 becomes the effect of x_1 when x_2 is at its mean (not 0)
  - ✓ β_2 is the effect of x_2 when x_1 is 0

Graphical Illustration

When x_1 and x_2 are continuous and centered

$$y_c = \beta_1 x_{1c} + \beta_2 x_{2c} + \beta_{Int} x_{1c} x_{2c} + \varepsilon$$



R Tips

For continuous x continuous interaction effects, all variables involved in interaction terms and the response variable need to be centered first to avoid issues of high multicollinearity and scale invariance.

```
dataName$y=scale(dataName$y,center=TRUE,scale=FALSE)
```

Use a different name (e.g., `dataName$y.centered=etc`.) if you would like the centered values to be stored in different variables.

```
dataName$x1=scale(dataName$x1,center=TRUE,scale=FALSE)
```

```
dataName$x2=scale(dataName$x2,center=TRUE,scale=FALSE)
```

`lm(y~x1:x2,data=dataName)` → The main effects are unchanged, but the interaction effects are at the means of `x1` and `x2`, and they change as the values of `x1` and `x2` change.

Kogod School
of
Business