Additive Models vs Interaction models

the <u>effect</u> of a covariate is the impact that changing it has on the expected response.

Numeric covariate: effect is the slope parameter

factor covariate: effects are the difference in expected response from one level to another. (e.g. $b_2 - b_1$)

A model with multiple covariates is <u>additive</u> if the effect of each covariate does <u>not</u> depend on the values of the other variables

A model has an <u>interaction</u> between two variables if the effect of one variable depends on the value of the other.

By definition, models with interactions are not additive.

Example: additive model

factor: city (2 levels) $j(i)$

numeric: 7am temperature x_i

response: 3pm temperature y_i

model for y_i: $Y_i = b_0 + b_{j(i)} + c_0 x_i + \epsilon_i$

effect of 7am temp is c_0:

$$E(Y_1) = b_0 + b_1 + c_0 x_1$$
$$E(Y_2) = b_0 + b_1 + c_0 x_2$$
$$\longrightarrow E(Y_2) - E(Y_1) = c_0(x_2 - x_1)$$

effect of changing 7am temp by $x_2 - x_1$ is $c_0(x_2 - x_1)$

does not depend on city (same effect in both cities)

effect of city is $b_2 - b_1$:

$$E(Y_3) = b_0 + b_1 + c_0 47$$
$$E(Y_4) = b_0 + b_2 + c_0 47$$
$$\longrightarrow E(Y_4) - E(Y_3) = b_2 - b_1$$

effect of changing city while holding 7am temp constant is $b_2 - b_1$. Does not depend on value of 7am temp.

Example: interaction model

$$Y_i = b_0 + b_{j(i)} + (c_0 + c_{j(i)}) x_i + \varepsilon_i$$

$$E(Y_1) = b_0 + b_1 + (c_0 + c_1)52$$
$$E(Y_2) = b_0 + b_1 + (c_0 + c_1)54$$
$$\longrightarrow E(Y_2) - E(Y_1) = (c_0 + c_1)(2)$$

$$E(Y_5) = b_0 + b_2 + (c_0 + c_2)52$$
$$E(Y_6) = b_0 + b_2 + (c_0 + c_2)54$$
$$\longrightarrow E(Y_6) - E(Y_5) = (c_0 + c_2)(2)$$

in city 1, effect is $c_0 + c_1$

in city 2, effect is $c_0 + c_2$

$$E(Y_3) = b_0 + b_1 + (c_0 + c_1)47$$

$$E(Y_4) = b_0 + b_2 + (c_0 + c_2)47$$

$$E(Y_4) - E(Y_3) = b_2 - b_1 + (c_2 - c_1)47$$

effect of changing the city depends on 7am temp!

This example was a factor-numeric interaction model.

We can also have numeric-numeric interactions

and factor-factor interactions (2-factor interactions)

Factor-Numeric interactions

factor has J levels

$$Y_i = b_0 + b_{j(i)} + (c_0 + c_{j(i)})x_i + \varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

Each level gets its own regression line

intercept: $b_0 + b_k$

slope: $c_0 + c_k$

How many model degrees of freedom? $2J$

How many mean parameters? $2J + 2$

in R, default constraints are $b_1 = 0$ and $c_1 = 0$

<u>Exercise</u>: state the reduced model, the full model, the null hypothesis (in terms of parameters), and give the degrees of freedom in the F-test comparing the factor numeric additive model to the factor-numeric interaction model.

reduced model: $Y_i = b_0 + b_{j(i)} + c_0 x_i + \varepsilon_i$

full model: $Y_i = b_0 + b_{j(i)} + \left(c_0 + c_{j(i)}\right) x_i + \varepsilon_i$

$H_0 : c_1 = c_2 = \cdots = c_J$

numerator dof $= 2J - (J+1) = J-1$

denominator dof $= n - 2J$

Factor-factor interactions

Factor 1 : $j(i)$, J levels

Factor 2 : $k(i)$, K levels

$$Y_i = b_0 + b_{j(i)} + c_{k(i)} + (bc)_{j(i),\,k(i)} + \varepsilon_i$$

Example: $J=2$, $K=3$

$$\begin{array}{cc|c|c|c}
 & & & k & \\
 & & 1 & 2 & 3 \\
\hline
j & 1 & b_0 + b_1 + c_1 + (bc)_{11} & b_0 + b_1 + c_2 + (bc)_{12} & b_0 + b_1 + c_3 + (bc)_{13} \\
\hline
 & 2 & b_0 + b_2 + c_1 + (bc)_{21} & b_0 + b_2 + c_2 + (bc)_{22} & b_0 + b_2 + c_3 + (bc)_{23} \\
\hline
\end{array}$$

Each of the JK possible combinations gets its own unrestricted expected value.

model dof = JK, $1 + J + K + JK$ parameters

in R, default constraint sets every parameter with a "1" subscript equal to 0.

$$b_1 = c_1 = (bc)_{11} = \cdots = (bc)_{J1} = \cdots = (bc)_{1K} = 0$$

$$\begin{array}{cc|c|c|c}
 & & & k & \\
 & & 1 & 2 & 3 \\
\hline
j & 1 & b_0 & b_0 + c_2 & b_0 + c_3 \\
\hline
 & 2 & b_0 + b_2 & b_0 + b_2 + c_2 + (bc)_{22} & b_0 + b_2 + c_3 + (bc)_{23} \\
\hline
\end{array}$$

interactions:

i	j	k
1	1	1
2	1	2
3	2	1
4	2	2

$E(Y_2) - E(Y_1) = c_2$

$E(Y_4) - E(Y_3) = c_2 + (bc)_{22}$

$E(Y_3) - E(Y_1) = b_2$

$E(Y_4) - E(Y_2) = b_2 + (bc)_{22}$

No interaction means

$$\big(E(Y_4) - E(Y_2)\big) - \big(E(Y_3) - E(Y_1)\big) = (bc)_{22} = 0$$

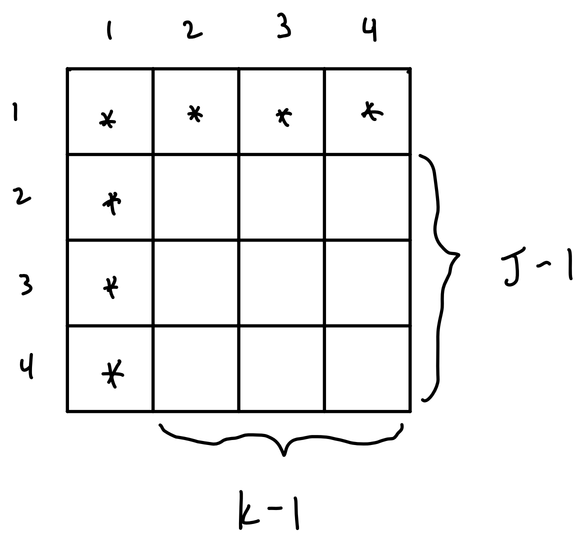
"difference of differences"

Degrees of freedom:

Interaction model: JK

Additive Model: $J + K - 1$

Difference $= JK - J - K + 1 = (J-1)(K-1)$



Numeric – numeric interactions

x_{i1} – numeric covariate

x_{i2} – numeric covariate

y_i – response

Model: $Y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + b_3 x_{i1} x_{i2} + \varepsilon_i$

$$\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

This is a multiple linear model $\longrightarrow$ define $x_{i3} = x_{i1} x_{i2}$

example $X = \begin{bmatrix} 1 & 3 & 5 & 15 \\ 1 & 4 & 1 & 4 \\ 1 & 7 & 8 & 56 \\ 1 & 2 & 12 & 24 \\ 1 & 6 & 1 & 6 \end{bmatrix}$

The effect of x_{i1} depends on x_{i2}:

$$E(Y_i) = (b_0 + b_2 x_{i2}) + (b_1 + b_3 x_{i2}) x_{i1}$$

intercept : $b_0 + b_2 x_{i2}$

slope : $b_1 + b_3 x_{i2}$ $\leftarrow$ effect depends on x_{i2}

The effect of x_{i2} depends on x_{i1}

$$E(Y_i) = (b_0 + b_1 x_{i1}) + (b_2 + b_3 x_{i1}) x_{i2}$$

intercept : $b_0 + b_1 x_{i1}$

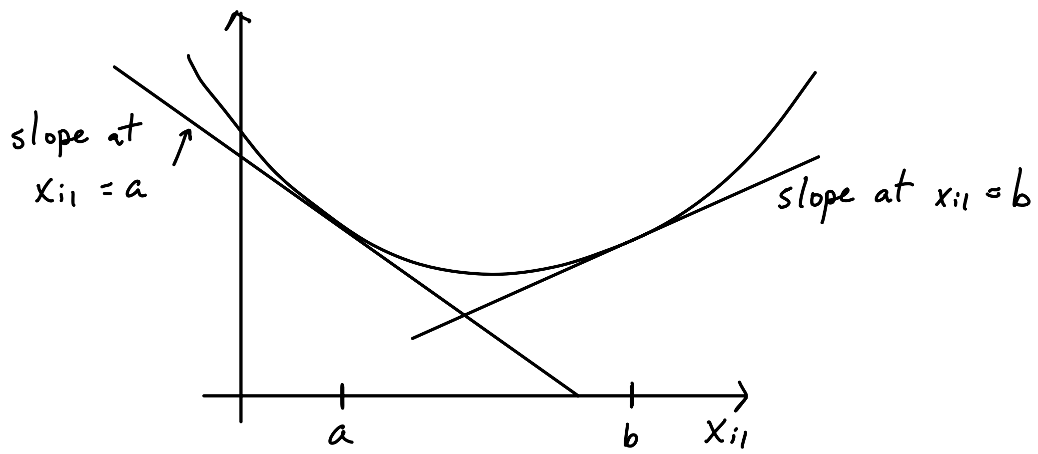
slope : $b_1 + b_3 x_{i1}$ $\leftarrow$ effect depends on x_{i1}

Quadratic Models

$$Y_i = b_0 + b_1 x_{i1} + b_2 x_{i1}^2 + \varepsilon_i$$

this is still a multiple linear model: $x_{i2} = x_{i1}^2$

Can think of this as an interaction btw x_{i1} and x_{i1}



Interpretations

$b_0 = E(Y_i)$ when $x_{i1} = 0$

$$\frac{d\,E(Y_i)}{d x_{i1}} = b_1 + 2 b_2 x_{i1}$$ effect (slope) depends on x_{i1}

b_1 = slope at $x_{i1} = 0$

b_2 tells us whether $E(Y_i)$ concave up or down.

$\frac{-b_1}{2b_2}$ = location of maximum or minimum.

polynomial models:

$$Y_i = \sum_{j=0}^{p} b_j x_i^j + \varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

Two-variable quadratic models

$$Y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + b_{12} x_{i1} x_{i2} + b_{11} x_{i1}^2 + b_{22} x_{i2}^2 + \varepsilon_i$$

$$= (b_0 + b_2 x_{i2} + b_{22} x_{i2}^2) + (b_1 + b_{12} x_{i2}) x_{i1} + (b_{11}) x_{i1}^2 + \varepsilon_i$$

$$= (b_0 + b_1 x_{i1} + b_{11} x_{i1}^2) + (b_2 + b_{12} x_{i1}) x_{i2} + (b_{22}) x_{i2}^2 + \varepsilon_i$$

additive in x_{i1} and x_{i2} if $b_{12} = 0$

$$Y_i = b_0 + [x_{i1} \ x_{i2}] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \frac{1}{2} [x_{i1} \ x_{i2}] \begin{bmatrix} 2b_{11} & b_{12} \\ b_{12} & 2b_{22} \end{bmatrix} \begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix} + \varepsilon_i$$

can add more predictors and have quadratic model in all variables

$$Y_i = b_0 + \underline{x}_i^T \underline{b} + \frac{1}{2} \underline{x}_i^T B \underline{x}_i + \varepsilon_i \qquad [B]_{ij} = b_{ij}$$

$$= b_0 + \sum_{j=1}^{p} x_{ij} b_j + \frac{1}{2} \sum_{j=1}^{p} \sum_{k=1}^{p} x_{ij} x_{ik} b_{jk} + \varepsilon_i$$

Useful in situations where you want to find an optimal $\underline{x}$ for some process.

if B is p.d. $\underline{x}^* =$ value that minimizes $E(Y_i)$

$$\nabla_x E(Y_i) = \underline{b} + B\underline{x}$$

$$\Rightarrow \underline{x}^* = -B^{-1}\underline{b}$$