V. Multiple Regression

0

5.1 Multiple covariates in a linear model

1

Data $(Y_i, Z_i^T)^T$, $i = 1, \dots, n$

$\sim (Y, Z^T)^T$ (generic random vector)

$Z \in \mathcal{Z} \subseteq \mathbb{R}^p$, $Z = (z_1, \dots, z_p)^T$

AIM: Specify a (linear) model for $E(Y | Z = z)$
$=: m(z)$ regress. function

LINEAR MODEL: $m(z) = (t(z))^T \beta$

Usual approach: $m(z)$ is based on additive and interaction terms based on parameterizations $s_1, \dots, s_p$ of single covariates $z_1, \dots, z_p$.

5.1.1 Additivity

Def 9.1 Additivity of the covariate effect

We say that a covariate ~~Z~~ Z_1 acts additively in the regression model with covariates $Z = (Z_1, \dots, Z_p)^T \in \mathcal{Z} \subseteq \mathbb{R}^p$ if the regression function is of the form

$$E(Y \mid Z_1 = z_1, Z_2 = z_2, \dots, Z_p = z_p) = m_1(z_1) + m_2(z_{(-1)}),$$

where $z_{(-1)} = (z_2, \dots, z_p)^T$, $m_1 : \mathbb{R} \to \mathbb{R}$ and $m_2 : \mathbb{R}^{p-1} \to \mathbb{R}$ are some measurable functions.

Example: Z_1 parameterized by $\delta_1(z_1)$ and the regression function is

$$E(Y \mid Z = z) = \beta_0 + \delta_1^T(z_1)\beta^1 + \underbrace{m_2(z_{(-1)})}_{\text{whatever}}$$

It is implied by additivity:

$\forall z_1 \in \mathbb{R},\ z_{(-1)} \in \mathbb{R}^{p-1}$

$$\mathsf{E}(Y \mid Z_1 = z_1 + \delta,\ Z_{(-1)} = z_{(-1)}) - \mathsf{E}(Y \mid Z_1 = z_1,\ Z_{(-1)} = z_{(-1)})$$
$$= m_1(z_1 + \delta) - m_1(z_1)$$

That is, the influence (effect) of the covariate Z_1 on the response expectation is the same for any value (does not depend on a value) of the remaining covariates $Z_{(-1)}$.

Terminology:

Quantity $m_1(z_1 + 1) - m_1(z_1)$ is called partial effect of the covariate Z_1 on the response expectation (effect adjusted for[CZ] possible influence of the remaining covariates $Z_{(-1)} = (Z_2, \dots, Z_p)^T$).

CZ: očištěný od možného vlivu.

Example: Linear model with regressor = covariate for effect of Z_1.

$$\mathsf{E}(Y \mid Z_1 = z_1,\ Z_{(-1)} = z_{(-1)}) = \beta_0 + \beta_1 z_1 + \underset{\text{whatever}}{m_2(z_{(-1)})}$$

$\Rightarrow \forall z_1 \in \mathbb{R},\ z_{(-1)} \in \mathbb{R}^{p-1}$

$$\mathsf{E}(Y \mid Z_1 = z_1 + 1,\ Z_{(-1)} = z_{(-1)}) - \mathsf{E}(Y \mid Z_1 = z_1,\ Z_{(-1)} = z_{(-1)})$$
$$= \beta_1$$

$\boxed{2}$

Remark: Z_1 parameterized by $\mathbb{S}_1(z_1)$ and the regression function is

$$E(Y|Z=z) = \beta_0 + \mathbb{S}_1^T(z_1)\beta + m_2(z_{(-1)})$$

then with data $(Y_i, Z_i^T)^T$, $i=1,\dots,n$ the model matrix is

$$X = \begin{pmatrix} 1 & \mathbb{S}_1^T(z_{1,1}) & * & \cdots & * \\ \vdots & \vdots & & & \\ 1 & \mathbb{S}_1^T(z_{n,1}) & * & \cdots & * \end{pmatrix}$$

$\mathbb{S}_1$: reparametrizing matrix of covariate Z_1

5.1.2 Interactions

Interactions are one possible way on how to model a situation that effect of some covariates is not additive.

APPENDIX

Def A.4 Elementwise product of two vectors

Let $a = (a_1, \dots, a_p)^T \in \mathbb{R}^p$, $c = (c_1, \dots, c_p)^T \in \mathbb{R}^p$. Their elementwise product is a vector $(a_1 \cdot c_1, \dots, a_p \cdot c_p)^T$ that will be denoted as $a : c$. That is,

$$a : c = \begin{pmatrix} a_1 \cdot c_1 \\ \vdots \\ a_p \cdot c_p \end{pmatrix}$$

$\equiv$ operator $*$ if applied to vectors in R

Def A.5 Columnwise product of two matrices

Let $A_{n \times p} = (a^1, \dots, a^p)$ and $C_{n \times q} = (c^1, \dots, c^q)$ be real matrices. Their columnwise product $A : C$ is a matrix $D_{n \times p \cdot q}$ such that

$$D = A : C := (a^1 : c^1, \dots, a^p : c^1, \dots, a^1 : c^q, \dots, a^p : c^q).$$

3

For specialists only

$$A = (a^1, \dots, a^p) = \begin{pmatrix} a_1^T \\ \vdots \\ a_n^T \end{pmatrix} \leftarrow \text{rows}$$

$$C = (c^1, \dots, c^q) = \begin{pmatrix} c_1^T \\ \vdots \\ c_n^T \end{pmatrix}$$

columns

$$A : C = (a^1 : c^1, \dots, a^p : c^1, \dots\dots, a^1 : c^q, \dots, a^p : c^q)$$

$$= \begin{pmatrix} c_1^T \otimes a_1^T \\ \vdots \\ c_n^T \otimes a_n^T \end{pmatrix}$$

← useful if we want to calculate rank$(A:C)$ from rank(A), rank(C)

POSSIBLE ALTERNATIVE (SENSIBLE) DEFINITION OF $A:C$ could be

$$A : C = \begin{pmatrix} a_1^T \otimes c_1^T \\ \vdots \\ a_n^T \otimes c_n^T \end{pmatrix} = (a^1 : c^1, \dots, a^1 : c^q, \dots$$

$$\dots a^p : c^1, \dots, a^p : c^q)$$

Def 5.2 Interaction terms

Let $(Z,W)^T \in \mathcal{Z} \times \mathcal{W} \subseteq \mathbb{R}^2$ be two covariates being parameterized using parameterizations

$\mathbb{s}_Z : \mathcal{Z} \to \mathbb{R}^{k-1}$ $\left(\mathbb{s}_Z = (s_Z^1, \dots, s_Z^{k-1})^T\right)$ and

$\mathbb{s}_W : \mathcal{W} \to \mathbb{R}^{l-1}$ $\left(\mathbb{s}_W = (s_W^1, \dots, s_W^{l-1})^T\right)$.

By interaction terms based on those two parameterizations we mean elements of a vector

$$\mathbb{s}_{ZW}(Z,W) := \mathbb{s}_W^T(W) \otimes \mathbb{s}_Z^T(Z)$$

$$= \left(s_Z^1(Z)\cdot s_W^1(W), \dots, s_Z^{k-1}(Z)\cdot s_W^1(W), \dots \right.$$

$$\left. \dots s_Z^1(Z)\cdot s_W^{l-1}(W), \dots, s_Z^{k-1}(Z)\cdot s_W^{l-1}(W)\right)^T$$

Usage of interaction terms in a linear model

Suppose that data are $(Y_i, Z_i, W_i)^T \sim (Y, Z, W)$, $i = 1, \dots, n$, ~~and~~ Z parameterized using $\mathbb{s}_Z$, W parameterized using $\mathbb{s}_W$.

Model which assumes **additivity** could be:

$$M_{Z+W}: \quad \boxed{\mathbb{E}(Y \mid Z = z, W = w) = \beta_0 + \mathbb{s}_Z^T(z)\underbrace{\beta^Z}_{(\beta^Z_1, \dots, \beta^Z_{k-1})^T} + \mathbb{s}_W^T(w)\underbrace{\beta^W}_{(\beta^W_1, \dots, \beta^W_{l-1})^T}}$$

the model matrix with data $(Y_i, Z_i, W_i)^T$, $i = 1, \dots, n$

$$X_{Z+W} = \begin{pmatrix} 1 & \mathcal{S}_Z^T(Z_1) & \mathcal{S}_W^T(W_1) \\ \vdots & \vdots & \vdots \\ 1 & \mathcal{S}_Z^T(Z_n) & \mathcal{S}_W^T(W_n) \end{pmatrix}$$

$\underbrace{\qquad}_{\mathcal{S}_Z}$ $\underbrace{\qquad}_{\mathcal{S}_W}$ (reparameterizing matrices for Z and W, respectively

num. of cols: 1 $\quad k-1 \quad l-1$

- X_{Z+W} will be of full-rank if
  - $\mathbb{1} \notin \mathcal{M}(\mathcal{S}_Z)$, $\mathbb{1} \notin \mathcal{M}(\mathcal{S}_W)$
  - columns in $\mathcal{S}_Z$ linearly indep. with columns in $\mathcal{S}_W$ $\left(\operatorname{rank}(\mathcal{S}_Z, \mathcal{S}_W) = k+l-2\right)$

Model which no more assumes additivity and allows for so called <u>effect modification</u> would be

$$M_{ZW}: \ \mathbb{E}(Y | Z=z, W=w) =$$

$$= \beta_0 + \mathcal{S}_Z^T(z)\underbrace{\beta^Z}_{(\beta_1^Z, \dots, \beta_{k-1}^Z)^T} + \mathcal{S}_W^T(w)\underbrace{\beta^W}_{(\beta_1^W, \dots, \beta_{\ell-1}^W)^T} + \mathcal{S}_{ZW}^T(z,w)\underbrace{\beta^{ZW}}_{(\beta_{11}^{ZW}, \dots, \beta_{k-1,1}^{ZW}, \dots \beta_{1,\ell-1}^{ZW}, \dots, \beta_{k-1,\ell-1}^{ZW})^T}$$

$\to$ the model matrix would become

$$X_{ZW} = \begin{pmatrix} 1 & \mathcal{S}_Z^T(Z_1) & \mathcal{S}_W^T(W_1) & \mathcal{S}_W^T(W_1) \otimes \mathcal{S}_Z^T(Z_1) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \mathcal{S}_Z^T(Z_n) & \mathcal{S}_W^T(W_n) & \mathcal{S}_W^T(W_n) \otimes \mathcal{S}_Z^T(Z_n) \end{pmatrix} =$$

$$= \begin{array}{cccc} \big(\mathbb{1} & \mathcal{S}_Z & \mathcal{S}_W & \mathcal{S}_Z : \mathcal{S}_W \big) \\ 1 & k-1 & \ell-1 & (k-1)\cdot(\ell-1) \end{array}$$

Can be shown:

If $\mathbb{1} \notin \mathcal{M}(S_Z)$, $\mathbb{1} \notin \mathcal{M}(S_W)$,
$\operatorname{rank}(S_Z, S_W) = k + l - 2$

$$\Rightarrow \quad \operatorname{rank}(S_Z : S_W) = (k-1)\cdot(l-1)$$

$$\operatorname{rank}(\mathbb{1}, S_Z, S_W, S_Z : S_W) =$$

$$= 1 + (k-1) + (l-1) + (k-1)(l-1)$$

$$= k \cdot l$$ (= number of columns of the X_{ZW} matrix)

Will follow:

meaning and interpretation of models being built using above principles when

- Z categorical, W numeric
- both Z, W numeric
- both Z, W categorical