Centered Explanatory Variables[1]

STA302 Fall 2020

[1] See last slide for copyright information.

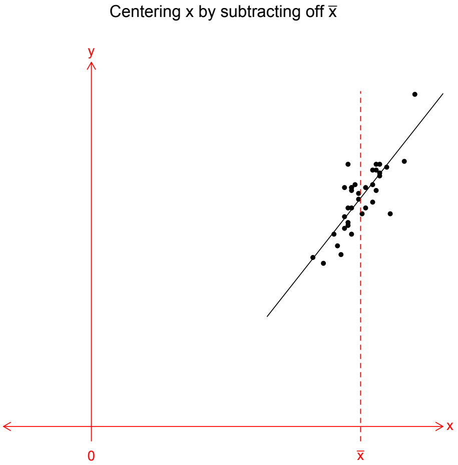
Overview

Center the explanatory variables

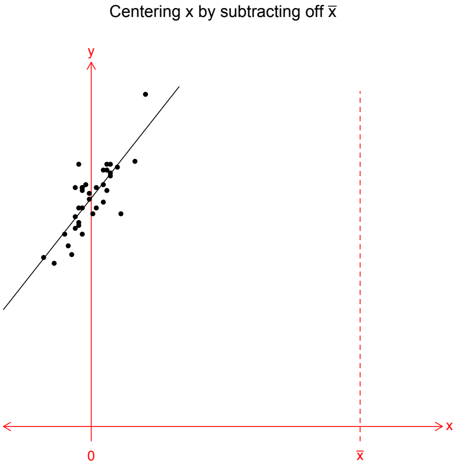
By subtracting off the sample mean

- Replace x_{ij} with $x_{ij} - \overline{x}_j$, expressing each explanatory variable as a deviation from its mean.
- Can be useful at times.

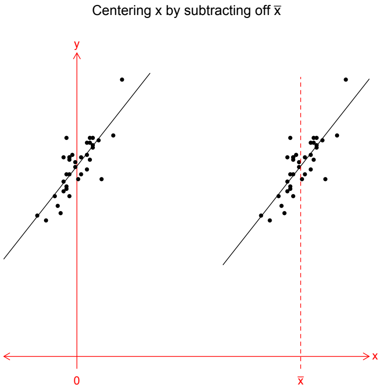
Simple Regression



Simple Regression

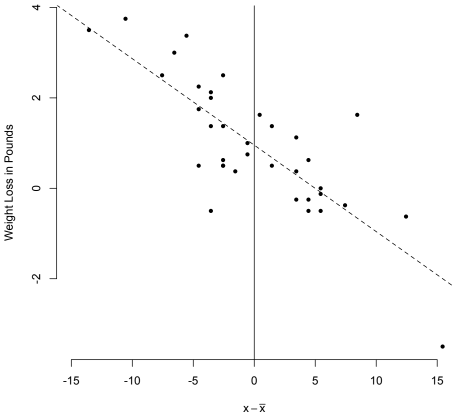


It looks like



- Estimated slopes will be unaffected.
- Estimated intercepts *will* be affected.
- $\widehat{y}_i$ should be unaffected.
- $\widehat{\epsilon}_i$ should be unaffected.
- If so, prediction intervals and R^2 should be unaffected.
- And tests for slopes should be unaffected.

Interpretation



- Having the y axis go through the data can make the intercept more meaningful.
- Suppose x is age, and y is weight loss in an exercise program.
- Question: Is any weight loss to be expected for a person of average age?
- $H_0 : \beta_0 = 0$ is tested automatically.
- Testing $H_0 : \beta_0 + \beta_1 \overline{x} = 0$ requires more effort.

The Model for Simple Regression

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_i + \epsilon_i \\
&= \beta_0 + \beta_1 x_i - \beta_1 \overline{x} + \beta_1 \overline{x} + \epsilon_i \\
&= (\beta_0 + \beta_1 \overline{x}) + \beta_1 (x_i - \overline{x}) + \epsilon_i \\
&= \quad \alpha_0 \quad + \alpha_1 (x_i - \overline{x}) + \epsilon_i
\end{aligned}
$$

The intercept is affected by centering, but the slope is not.

Center all the predictor variables

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k} + \epsilon_i \\
&= \beta_0 + \beta_1 \overline{x}_1 + \cdots + \beta_k \overline{x}_k \\
&\quad + \beta_1 (x_{i,1} - \overline{x}_1) + \cdots + \beta_k (x_{i,k} - \overline{x}_k) + \epsilon_i \\
&= \alpha_0 + \alpha_1 (x_{i,1} - \overline{x}_1) + \cdots + \alpha_k (x_{i,k} - \overline{x}_k) + \epsilon_i
\end{aligned}
$$

with

$$\alpha_0 = \beta_0 + \beta_1 \overline{x}_1 + \cdots + \beta_k \overline{x}_k.$$

$$\alpha_j = \beta_j \text{ for } j = 1, \ldots, k$$

- The intercept is affected by centering, but the slopes are not.
- You don't have to center all the x variables.

Dummy Variable Regression

Just center the covariate(s)

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_i + \beta_2 d_{i,1} + \beta_3 d_{i,2} + \epsilon_i \\
&= \beta_0 + \beta_1 x_i - \beta_1 \overline{x} + \beta_1 \overline{x} + \beta_2 d_{i,1} + \beta_3 d_{i,2} + \epsilon_i \\
&= (\beta_0 + \beta_1 \overline{x}) + \beta_1 (x_i - \overline{x}) + \beta_2 d_{i,1} + \beta_3 d_{i,2} + \epsilon_i \\
&= \alpha_0 + \alpha_1 (x_i - \overline{x}) + \alpha_2 d_{i,1} + \alpha_3 d_{i,2} + \epsilon_i
\end{aligned}
$$

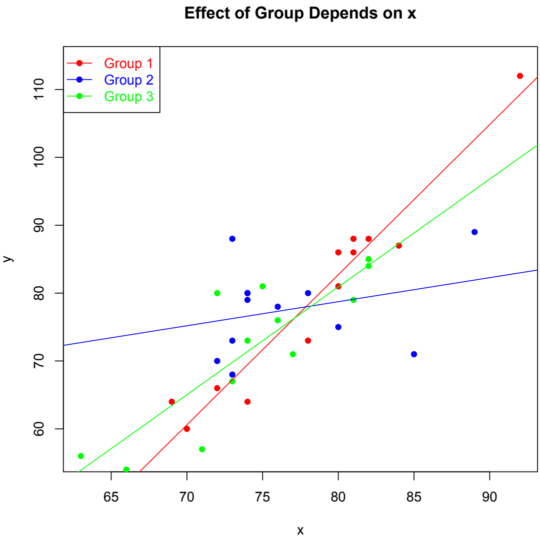
Slopes are not affected by centering.

Parallel Regression Lines

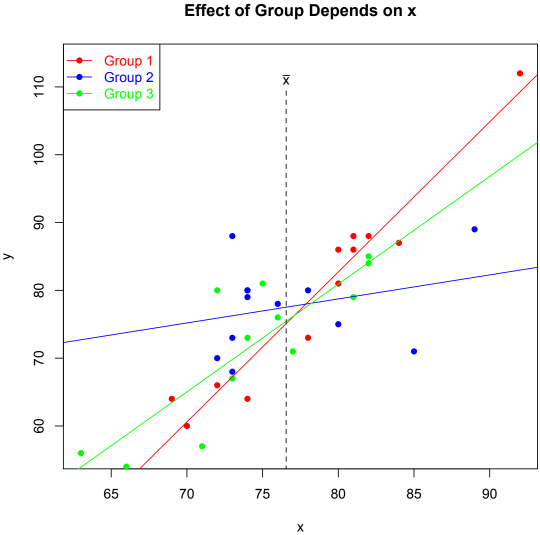
Drug	d_1	d_2	$E(y\|\mathbf{x}) = \alpha_0 + \alpha_1(x - \overline{x}) + \alpha_2 d_1 + \alpha_3 d_2$
A	1	0	$(\alpha_0 + \alpha_2) + \alpha_1(x - \overline{x})$
B	0	1	$(\alpha_0 + \alpha_3) + \alpha_1(x - \overline{x})$
Placebo	0	0	$\alpha_0 \qquad + \alpha_1(x - \overline{x})$

Could describe the estimated intercepts as “adjusted means,” or “corrected means.”

Interactions



Interactions



Interactions

Group	d_1	d_2	$E(y\|\mathbf{x})$
1	1	0	$(\beta_0 + \beta_2) + (\beta_1 + \beta_4)(x - \overline{x})$
2	0	1	$(\beta_0 + \beta_3) + (\beta_1 + \beta_5)(x - \overline{x})$
3	0	0	$\beta_0 + \beta_1 (x - \overline{x})$

- What happens at $x = \overline{x}$?
- If you are interested in estimating or testing for differences at some other point, it might be easiest to subtract that value from x instead.

Estimation and Testing

- Have
$$\alpha_0 = \beta_0 + \beta_1 \overline{x}_1 + \cdots + \beta_k \overline{x}_k.$$
$$\alpha_j = \beta_j \text{ for } j = 1, \ldots, k$$
- Will have
$$\widehat{\alpha}_0 = \widehat{\beta}_0 + \widehat{\beta}_1 \overline{x}_1 + \cdots + \widehat{\beta}_k \overline{x}_k.$$
$$\widehat{\alpha}_j = \widehat{\beta}_j \text{ for } j = 1, \ldots, k$$
- $\widehat{\mathbf{y}}$ will be unaffected.
- $\widehat{\boldsymbol{\epsilon}}$ will be unaffected.
- Prediction intervals and R^2 will be unaffected.
- Tests for slopes will be unaffected.

Re-parameterization

- The mapping

  $$\alpha_0 = \beta_0 + \beta_1\overline{x}_1 + \cdots + \beta_k\overline{x}_k.$$
  $$\alpha_j = \beta_j \text{ for } j = 1, \ldots, k$$

  is a one-to-one re-parameterization.
- Furthermore, it's linear.
- Write as matrix multiplication.

Matrix Multiplication

To get $\alpha_0 = \beta_0 + \beta_1\overline{x}_1 + \cdots + \beta_k\overline{x}_k$ and $\alpha_j = \beta_j$ for $j = 1, \ldots, k$

$$\begin{pmatrix} 1 & \overline{x}_1 & \overline{x}_2 & \cdots & \overline{x}_k \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} \beta_0 + \beta_1\overline{x}_1 + \cdots + \beta_k\overline{x}_k \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}$$

This matrix $\uparrow$ definitely has an inverse.

Inverse

$$\begin{pmatrix} 1 & -\overline{x}_1 & -\overline{x}_2 & \cdots & -\overline{x}_k \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & \overline{x}_1 & \overline{x}_2 & \cdots & \overline{x}_k \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$\mathbf{A} \qquad \mathbf{A}^{-1} \qquad = \qquad \mathbf{I}$$

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \\ &= \mathbf{X}\mathbf{A}\mathbf{A}^{-1}\boldsymbol{\beta} + \boldsymbol{\epsilon} \\ &= (\mathbf{X}\mathbf{A})(\mathbf{A}^{-1}\boldsymbol{\beta}) + \boldsymbol{\epsilon} \\ &= \quad \mathbf{W} \quad\quad \boldsymbol{\alpha} \quad\quad + \boldsymbol{\epsilon}, \end{aligned}$$

Where $\mathbf{W}$ is the centered $\mathbf{X}$ matrix.

Does the matrix $\mathbf{A}$ really center the $\mathbf{X}$ matrix?

Just look at row i of $\mathbf{XA}$

$$\left(\begin{array}{ccccc} 1 & x_{i1} & x_{i2} & \cdots & x_{ik} \end{array} \right) \left(\begin{array}{ccccc} 1 & -\overline{x}_1 & -\overline{x}_2 & \cdots & -\overline{x}_k \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{array} \right)$$

$$= \left(\begin{array}{ccccc} 1 & x_{i1} - \overline{x}_1 & x_{i2} - \overline{x}_2 & \cdots & x_{ik} - \overline{x}_k \end{array} \right)$$

One-to-one linear transformation

The point is that centering the explanatory variables is a one-to-one linear transformation of $\mathbf{X}$ matrix: $\mathbf{W} = \mathbf{AX}$.

$$\mathbf{A} = \begin{pmatrix} 1 & -\overline{x}_1 & -\overline{x}_2 & \cdots & -\overline{x}_k \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Centering Just Some of the Variables

$$\begin{pmatrix} 1 & x_{i1} & x_{i2} & \cdots & x_{ik} \end{pmatrix} \begin{pmatrix} 1 & -\overline{x}_1 & -\overline{x}_2 & \cdots & -\overline{x}_k \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

- To leave variable j uncentered, replace $\overline{x}_j$ with zero.
- Rows are still linearly independent.

We have been here before

See Assignment 9, Problem 4

$$
\begin{aligned}
& \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \\
\iff \quad & \mathbf{y} = \mathbf{X}\mathbf{A}\mathbf{A}^{-1}\boldsymbol{\beta} + \boldsymbol{\epsilon} \\
\iff \quad & \mathbf{y} = \mathbf{W}\boldsymbol{\alpha} + \boldsymbol{\epsilon}
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\boldsymbol{\alpha}} &= (\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{y} \\
&= \left((\mathbf{X}\mathbf{A})'\mathbf{X}\mathbf{A}\right)^{-1}(\mathbf{X}\mathbf{A})'\mathbf{y} \\
&= \left(\mathbf{A}'\mathbf{X}'\mathbf{X}\mathbf{A}\right)^{-1}\mathbf{A}'\mathbf{X}'\mathbf{y} \\
&= \mathbf{A}^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'^{-1}\mathbf{A}'\mathbf{X}'\mathbf{y} \\
&= \mathbf{A}^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\
&= \mathbf{A}^{-1}\widehat{\boldsymbol{\beta}}
\end{aligned}
$$

$\widehat{\boldsymbol{\alpha}} = \mathbf{A}^{-1}\widehat{\boldsymbol{\beta}}$

Same form as $\boldsymbol{\alpha} = \mathbf{A}^{-1}\boldsymbol{\beta}$: Invariance

$$\begin{pmatrix} \widehat{\alpha}_0 \\ \widehat{\alpha}_1 \\ \widehat{\alpha}_2 \\ \vdots \\ \widehat{\alpha}_k \end{pmatrix} = \begin{pmatrix} 1 & \overline{x}_1 & \overline{x}_2 & \cdots & \overline{x}_k \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \vdots \\ \widehat{\beta}_k \end{pmatrix} = \begin{pmatrix} \widehat{\beta}_0 + \widehat{\beta}_1\overline{x}_1 + \cdots + \widehat{\beta}_k\overline{x}_k \\ \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \vdots \\ \widehat{\beta}_k \end{pmatrix}$$

Predicted $\mathbf{y}$ for $\mathbf{y} = \mathbf{W}\boldsymbol{\alpha} + \boldsymbol{\epsilon}$

$$
\begin{aligned}
\mathbf{W}\widehat{\boldsymbol{\alpha}} &= (\mathbf{XA})(\mathbf{A}^{-1}\widehat{\boldsymbol{\beta}}) \\
&= \mathbf{X}\widehat{\boldsymbol{\beta}} \\
&= \widehat{\mathbf{y}}
\end{aligned}
$$

- So $\widehat{\mathbf{y}}$ is unchanged by centering.
- This means $\widehat{\boldsymbol{\epsilon}}$, SSE, MSE and R^2 are also unchanged.

Prediction Intervals are unchanged

$\mathbf{x}_0'\widehat{\boldsymbol{\beta}} \pm t_{\alpha/2}\sqrt{MSE(1+\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)}$

The key is that you need to give it a vector of centered x variables: $\mathbf{x}_0^{*\prime} = \mathbf{x}_0'\mathbf{A} \iff \mathbf{x}_0^* = \mathbf{A}'\mathbf{x}_0$.

$$
\begin{aligned}
& \mathbf{x}_0^{*\prime}\widehat{\boldsymbol{\alpha}} \pm t_{\alpha/2}\sqrt{MSE(1+\mathbf{x}_0^{*\prime}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{x}_0^*)} \\
= \; & (\mathbf{x}_0'\mathbf{A})(\mathbf{A}^{-1}\widehat{\boldsymbol{\beta}}) \pm t_{\alpha/2}\sqrt{MSE(1+\mathbf{x}_0'\mathbf{A}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{A}'\mathbf{x}_0)} \\
= \; & \mathbf{x}_0'\widehat{\boldsymbol{\beta}} \pm t_{\alpha/2}\sqrt{MSE(1+\mathbf{x}_0'\mathbf{A}\left((\mathbf{X}\mathbf{A})'\mathbf{X}\mathbf{A}\right)^{-1}\mathbf{A}'\mathbf{x}_0)} \\
= \; & \mathbf{x}_0'\widehat{\boldsymbol{\beta}} \pm t_{\alpha/2}\sqrt{MSE(1+\mathbf{x}_0'\mathbf{A}\left(\mathbf{A}'\mathbf{X}'\mathbf{X}\mathbf{A}\right)^{-1}\mathbf{A}'\mathbf{x}_0)} \\
= \; & \mathbf{x}_0'\widehat{\boldsymbol{\beta}} \pm t_{\alpha/2}\sqrt{MSE(1+\mathbf{x}_0'\mathbf{A}\mathbf{A}^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'^{-1}\mathbf{A}'\mathbf{x}_0)} \\
= \; & \mathbf{x}_0'\widehat{\boldsymbol{\beta}} \pm t_{\alpha/2}\sqrt{MSE(1+\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)}
\end{aligned}
$$

Hypothesis tests: $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{t}$

Using $F^* = \frac{(\mathbf{C}\widehat{\boldsymbol{\beta}}-\mathbf{t})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}}-\mathbf{t})}{q\, MSE}$

$$
\begin{aligned}
\mathbf{C}\boldsymbol{\beta} = \mathbf{t} \quad & \Longleftrightarrow \quad (\mathbf{CA})(\mathbf{A}^{-1}\boldsymbol{\beta}) = \mathbf{t} \\
& \Longleftrightarrow \quad (\mathbf{CA})\boldsymbol{\alpha} = \mathbf{t}
\end{aligned}
$$

Look at the numerator of F^* for the centered data.

$$
\begin{aligned}
& (\mathbf{CA}\widehat{\boldsymbol{\alpha}} - \mathbf{t})'(\mathbf{CA}(\mathbf{W}'\mathbf{W})^{-1}(\mathbf{CA})')^{-1}(\mathbf{CA}\widehat{\boldsymbol{\alpha}} - \mathbf{t}) \\
= \; & (\mathbf{CAA}^{-1}\widehat{\boldsymbol{\beta}} - \mathbf{t})'(\mathbf{CA}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{A}'\mathbf{C}')^{-1}(\mathbf{CAA}^{-1}\widehat{\boldsymbol{\beta}} - \mathbf{t}) \\
= \; & (\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})'(\mathbf{CAA}^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'^{-1}\mathbf{A}'\mathbf{C}')^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t}) \\
= \; & (\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})
\end{aligned}
$$

- This is the numerator for the uncentered data, so the test statistics are equal.
- If the hypothesis does not involve α_0, you don't need to transform $\mathbf{C}$.

Summary

A simple story, in spite of all the technical details

- Centering some or all of the explanatory variables can be helpful.
- Only the intercept is affected.
- There is no effect on predicted y, residuals, R^2, or prediction intervals.
- There is no effect on tests and confidence intervals, unless the intercept is involved.

Copyright Information