

`lm()` uses the closed form equations to find the least squares estimates but there are other ways to do this optimization; one was discussed in the notes. We will have you go through this method here, comparing to the linear model you made in the problem set. Load the `flights` data into your R session.

The function `rss` below is meant to mirror the formula for RSS for a simple linear regression model:

$$\text{RSS}(b_0, b_1) = \sum_{i=1}^n (y_i - (b_0 + b_1 x_i))^2$$

```
rss <- function(coef) {
  sum((flights$avg_speed - (coef[1] + coef[2] * flights$distance))^2)
}
```

`coef[1]` and `coef[2]` represent  $b_0$  and  $b_1$ , respectively and refer to the first two elements of a vector called `coef` which is the argument of the function.

- Use the `optim()` template given to find the slope  $b_0$  and intercept  $b_1$  that minimizes the RSS function. You can do this by using the table below. In this table, we have given you two different sets of starting values for  $b_0$  and  $b_1$  to try:
  - $b_0 = 0, b_1 = 0$
  - Eyeball the slope and intercept from the scatterplot with the line superimposed on the data (from the problem set).

Write in your final values in each case in the spaces given. Comment on whether they match the results you got from your linear model object in the problem set.

```
optim(par = c( , ), fn = )
```

| Starting $b_0$                    | Starting $b_1$                | Final $b_0$ | Final $b_1$ |
|-----------------------------------|-------------------------------|-------------|-------------|
| 0                                 | 0                             |             |             |
| Eyeballed <i>intercept</i> : ____ | Eyeballed <i>slope</i> : ____ |             |             |

- Change the `rss` function from the last question so it captures the sum of the *absolute value of the residuals*. This is given by the function below.

$$\text{RSABS}(b_0, b_1) = \sum_{i=1}^n |y_i - (b_0 + b_1 x_i)|$$

Call this new function you make `rsabs`. You can fill in the template below.

```
rsabs <- function(coef) {
```

```
}
```

3. Use `rsabs` to estimate the parameters  $b_0$  and  $b_1$  using `optim()` as you did before for `rss`. Comment on whether your results matches the estimates that come out of `lm()` and from the `rss` function.

| Starting $b_0$                    | Starting $b_1$                | Final $b_0$ | Final $b_1$ |
|-----------------------------------|-------------------------------|-------------|-------------|
| 0                                 | 0                             |             |             |
| Eyeballed <i>intercept</i> : ____ | Eyeballed <i>slope</i> : ____ |             |             |