EEEB UN3005/GR5005
Lab - Week 10 - 01 and 03 April 2019
USE YOUR NAME HERE

Interactions

Exercise 1: Multiple Regression with an Interaction

Referencing the *Statistical Rethinking* text (see page 230 and relevant preceding code), reconstruct and fit the model `m7.9`. More specifically, this model uses the `tulips` dataset. It models `blooms` as the outcome variable of interest, with centered versions of `water` and `shade`, as well as the interaction between them, as continuous predictor variables.

After fitting the model, use `precis()` to display the 97% PIs for all model parameters. Double check to make sure your results generally align with what's shown in the book.

Finally, how do you interpret the intercept estimate in the context of this model?

```
data(tulips)
d = tulips
d$shade.c = d$shade - mean(d$shade)
d$water.c = d$water - mean(d$water)
m7.9 = map(
    alist(
        blooms ~ dnorm(mu, sigma),
        mu <- a +
                bW * water.c +
                bS * shade.c +
                bWS * water.c * shade.c,
        bW ~ dnorm(0, 100),
        bS ~ dnorm(0, 100),
        bWS ~ dnorm(0, 100),
        sigma ~ dunif(0, 100)),
    data = d,
    start = list(
        a = mean(d$blooms),
        bW = 0,
        bS = 0,
        bWS = 0,
        sigma = sd(d$blooms)))
precis(m7.9, prob = 0.97)

##          Mean StdDev   1.5%  98.5%
```

```
## a       128.99    8.70 110.11 147.88
## bW       74.96   10.60  51.95  97.97
## bS      -41.14   10.60 -64.14 -18.14
## bWS     -51.87   12.95 -79.97 -23.77
## sigma    45.22    6.15  31.87  58.58
```

Answer:

$$\mu = a + b_W W + b_S S + b_{WS} WS$$

The intercept a is the value of μ when both W and S is 0. Since both of these two variables is centered, a is the blooms value of the plants that have water value equals to `mean(water)` and shade value equals to `mean(shade)`.

Exercise 2: Triptych Plots for Both Predictor Variables

In the *Statistical Rethinking* book, Figure 7.7 (page 234), you'll see a series of "triptych" plots. Pay particular attention to the bottom row of the panel, which shows model-based predictions (from model m7.9) for the effect of `shade.c` on `blooms`. Because the model includes multiple predictor variables, it makes sense to plot this relationship across multiple values of the other predictor (`water.c`, hence the series of three plots in the bottom row), and because the model includes an interaction effect, the specific relationship between `shade.c` and `blooms` varies across the triptych plots.

To fully visualize the predictions from model m7.9, create two triptych plots. The first triptych plot should show the effect of `water.c` on `blooms`, plotted for three different values of `shade.c`. The second triptych plot should show the effect of `shade.c` on `blooms`, plotted for three different values of `water.c` (i.e., this second triptych plot should replicate what you see in the bottom row of the book's Figure 7.7). Clearly, modifying the book's code will help you in generating this output. Feel free to visualize the 97% interval of the mean using lines (as in the book code) or using a shaded interval (as has been demonstrated in multiple places throughout the book and lecture code using the `shade()` function). Also note, you can plot these two triptych plots together neatly with some modification to your plotting window. See the book's R code 7.28 box for a hint as to how...

Using the plots to help with your interpretation, at which value(s) of shade is the effect of water most extreme? At which value(s) of water is the effect of shade most extreme?

```
par(mfrow = c(2, 3))

shade.seq = -1:1
for (w in -1:1){
    dt = d[d$water.c == w,]
    plot(
```

```
        blooms ~ shade.c,
        data = dt,
        col = rangi2,
        main = paste("water.c = ", w),
        xaxp = c(-1, 1, 2),
        ylim = c(0, 362),
        xlab = "shade (centered)")
    mu = link(
        m7.9,
        data = data.frame(water.c = w, shade.c = shade.seq))
    mu.mean = apply(mu, 2, mean)
    mu.PI = apply(mu, 2, PI, prob = 0.97)
    lines(shade.seq, mu.mean)
    lines(shade.seq, mu.PI[1,], lty = 2)
    lines(shade.seq, mu.PI[2,], lty = 2)
}

water.seq = -1:1
for (s in -1:1){
    dt = d[d$shade.c == s,]
    plot(
        blooms ~ water.c,
        data = dt,
        col = rangi2,
        main = paste("shade.c = ", s),
        xaxp = c(-1, 1, 2),
        ylim = c(0, 362),
        xlab = "water (centered)")
    mu = link(
        m7.9,
        data = data.frame(water.c = water.seq, shade.c = s))
    mu.mean = apply(mu, 2, mean)
    mu.PI = apply(mu, 2, PI, prob = 0.97)
    lines(water.seq, mu.mean)
    lines(water.seq, mu.PI[1,], lty = 2)
    lines(water.seq, mu.PI[2,], lty = 2)
}
```
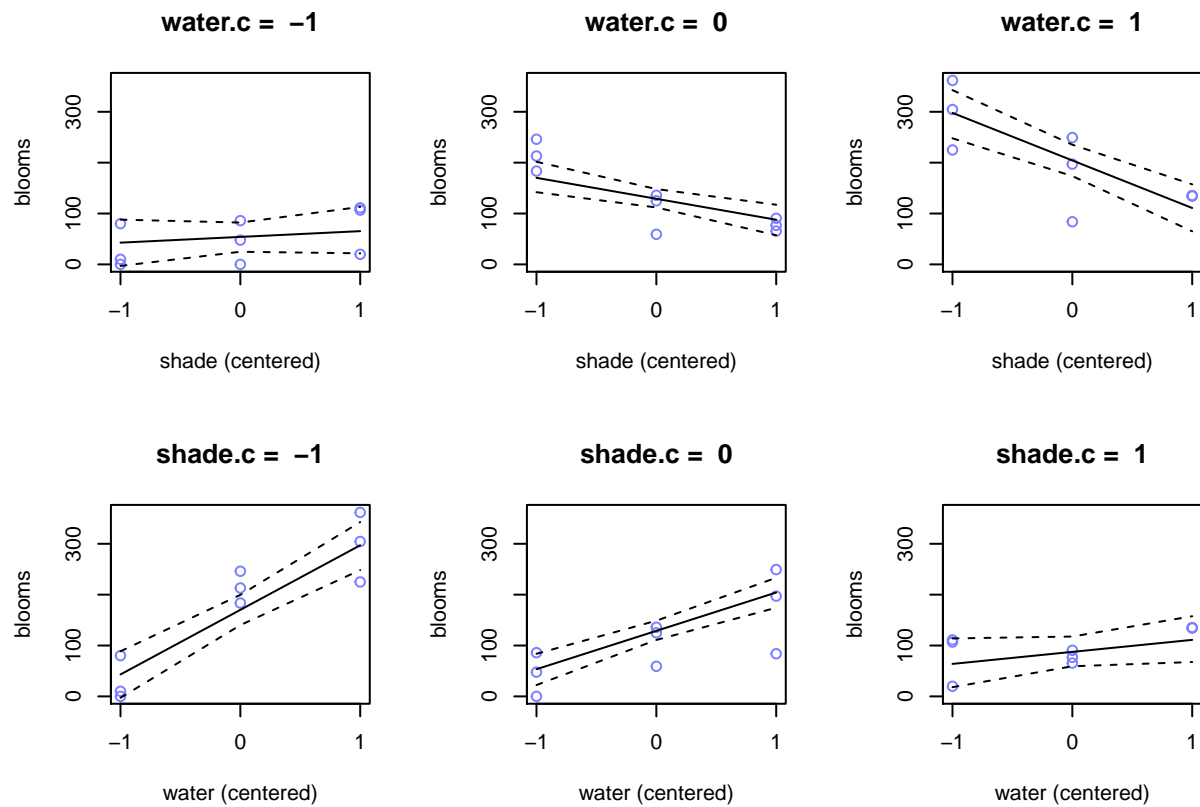
Answer:

With water value increasing, the blooms value is more and more negatively related to shade value.

Conversely, with shade value increasing, the blooms value is less and less positively related to water value.