This page shows how you can derive the simple form of the LSLR slope using only algebra:

$$b = \frac{s_y}{s_x} r$$

from its original definition.

$$b = \frac{\sum (x - \bar{x})(y - \bar{y})}{(x - \bar{x})^2}$$

Denominator looks a bit like variance:

$$s_x = \sqrt{\frac{\sum (x - \bar{x})^2}{(n-1)}}$$

$$s_x{}^2 = \frac{\sum (x - \bar{x})^2}{(n-1)}$$

$$s_x{}^2 (n-1) = \sum (x - \bar{x})^2$$

So we can rewrite b as:

$$b = \frac{\sum(x - \bar{x})(y - \bar{y})}{{s_x}^2(n - 1)}$$

And the factors of the numerator look a bit like z-scores…

$$z_x = \frac{(x - \bar{x})}{s_x}$$

$z_x s_x$ = $(x - \bar{x})$

$$z_y = \frac{(y - \bar{y})}{s_y}$$

$z_y s_y$ = $(y - \bar{y})$

So we can rewrite the numerator:

$\sum(x - \bar{x})(y - \bar{y})$ = $\sum z_x s_x z_y s_y$

And now we can rewrite b, using our simplified b from above:

$$b = \frac{\sum(x - \bar{x})(y - \bar{y})}{{s_x}^2(n-1)}$$

$$b = \frac{\sum z_x s_x z_y s_y}{{s_x}^2(n-1)}$$

But since standard deviation is a scalar (just a number, not a list of numbers), we can take it outside the summation:

$$b = \frac{s_x s_y \sum z_x z_y}{{s_x}^2(n-1)}$$

$$b = \frac{s_y \sum z_x z_y}{s_x(n-1)}$$

But this summation above should remind you of the simple form of Pearson's correlation formula!

$$r = \frac{\sum z_x z_y}{(n-1)}$$

So we can finally rewrite the slope of the LSLR line as:

$$b = \frac{s_y}{s_x} r$$

Note that if x has zero standard deviation, then b is undefined! But that's ok, because that means x is a constant, and therefore has no predictive power for y anyway.

In R code, b is given by:

```
lslr_slope = (sd(y)/sd(x))*cor(x, y)
```