

1 Some Important Definitions

To understand a diagnostic plot called the residual-leverage plot, we must understand three things:

- Leverage,
- Standardized residuals, and
- Cook's distance.

Example

Consider the plots associated with four different situations:

1. a dataset where everything is fine
2. a dataset with a high-leverage, but low-standardized residual point
3. a dataset with a low-leverage, but high-standardized residual point
4. a dataset with a high-leverage, high-standardized residual point

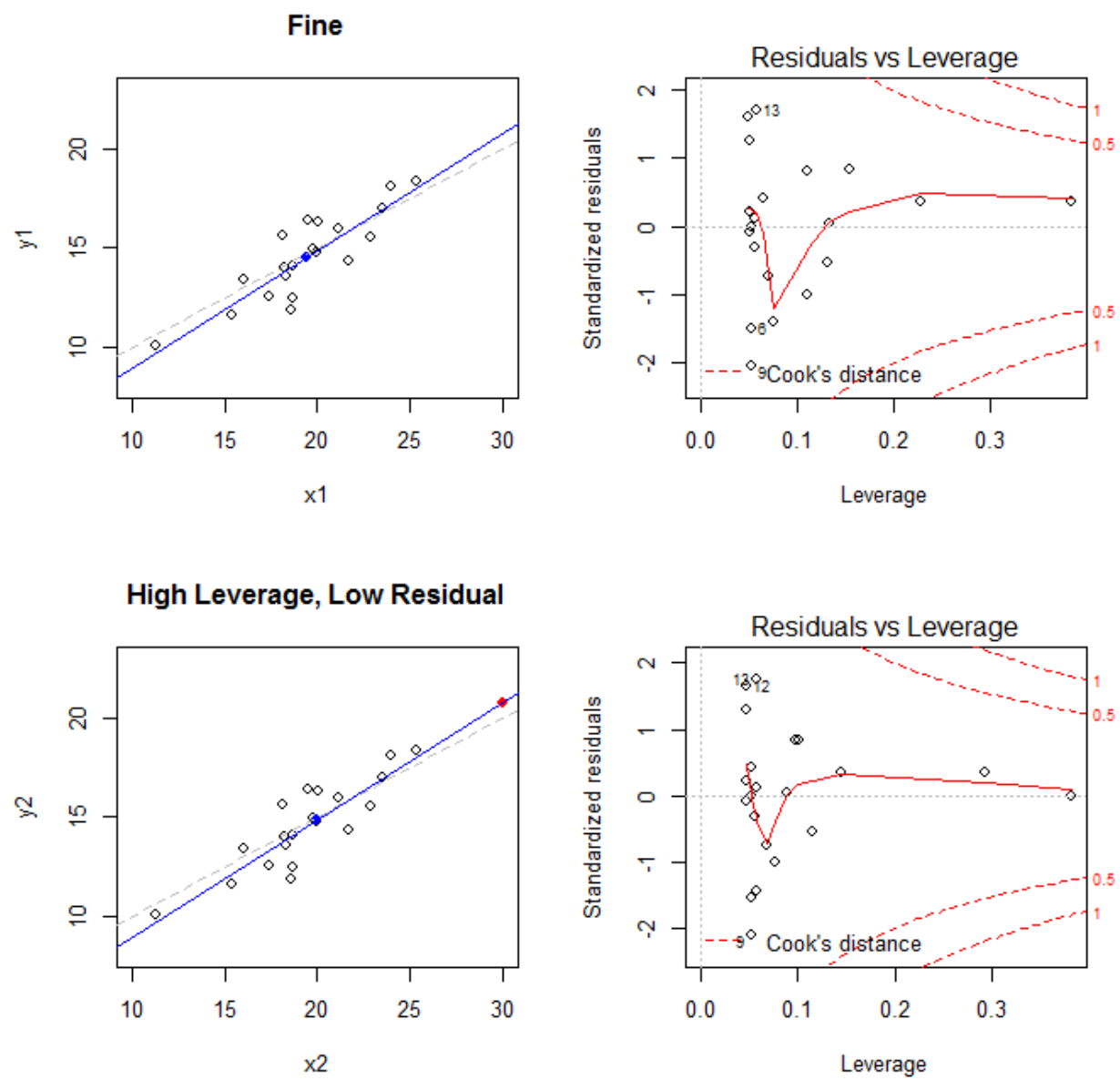


Figure 1:

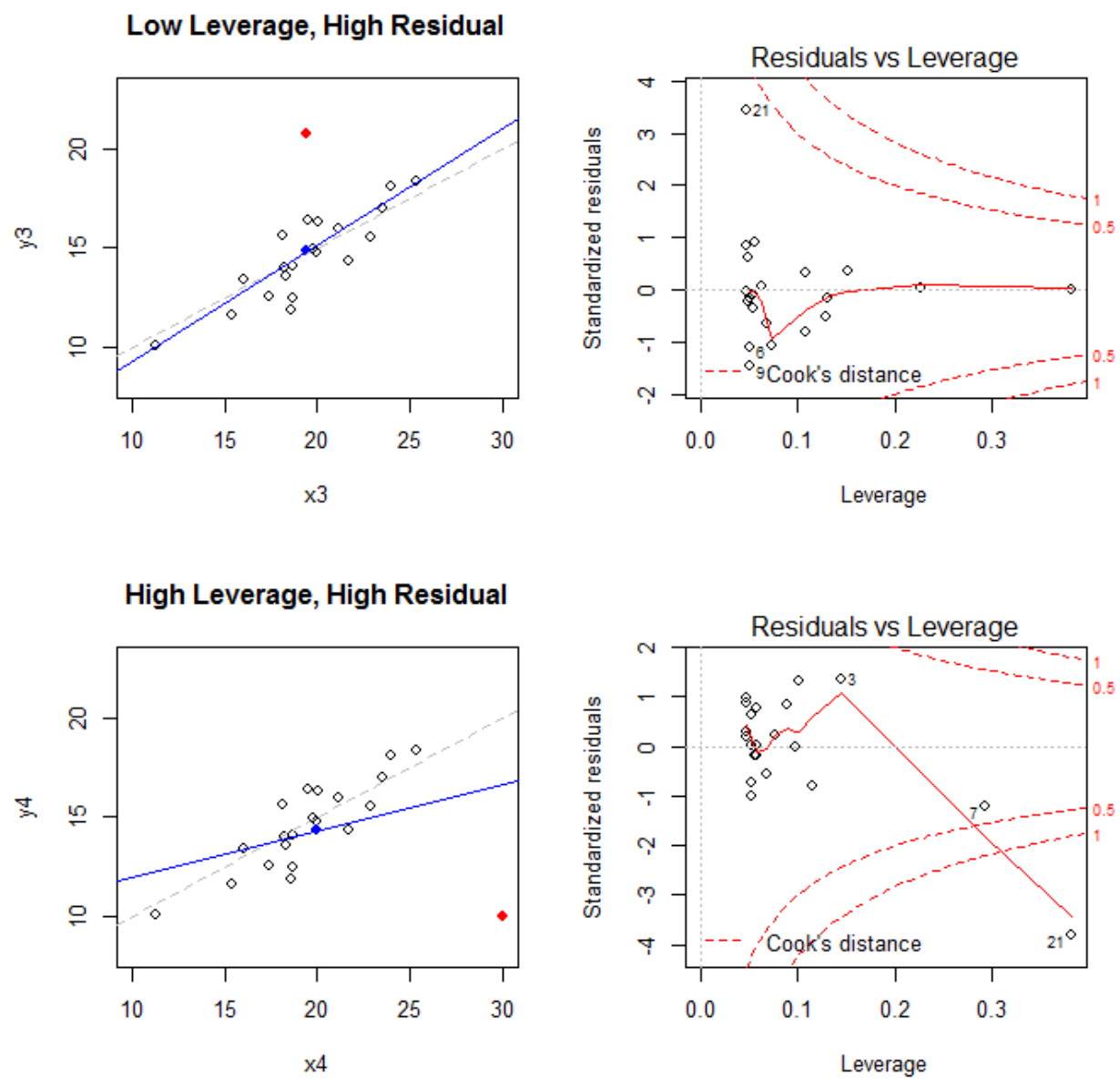


Figure 2:

- The plots on the left show the data, the center of the data with a blue dot, the underlying data generating process with a dashed gray line, the model fit with a blue line, and the special point with a red dot.
- On the right are the corresponding residual-leverage plots; the special point is 21.
- The model is badly distorted primarily in the fourth case where there is a point with high leverage and a large (negative) standardized residual.

For reference, here are the values associated with the special points:

	leverage	std.residual	cooks.d
high leverage, low residual	0.3814234	0.0014559	0.0000007
low leverage, high residual	0.0476191	3.4456341	0.2968102
high leverage, high residual	0.3814234	-3.8086475	4.4722437