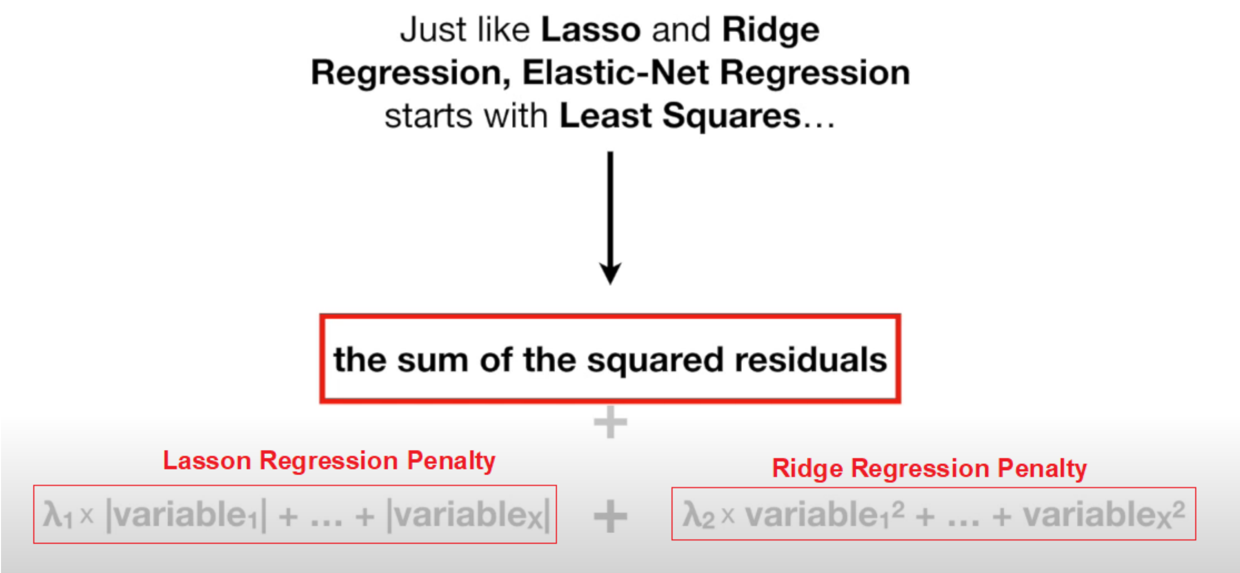
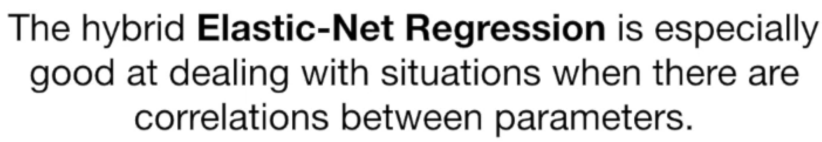
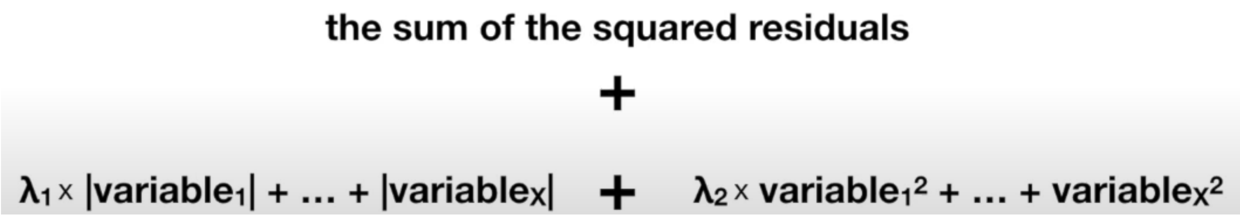
当参数有成千上万个时，可以使用Elastic Net Regression



注意，有两个不同的λ。当λ1等于0，λ2大于0时，就是ridge regression。反之为lasso regression。当λ1和λ2都等于0时，就是least suqare。

Elastic Net Regression的优点：





原因：

This is because on it's own, **Lasso Regression** tends to pick just one of the correlated terms and eliminates the others…

the sum of the squared residuals

+

$\lambda_1 \times |\text{variable}_1| + \ldots + |\text{variable}_X|$

…whereas **Ridge Regression** tends to shrink all of the parameters for the correlated variables together.

the sum of the squared residuals

+

$\lambda_2 \times \text{variable}_1^2 + \ldots + \text{variable}_X^2$