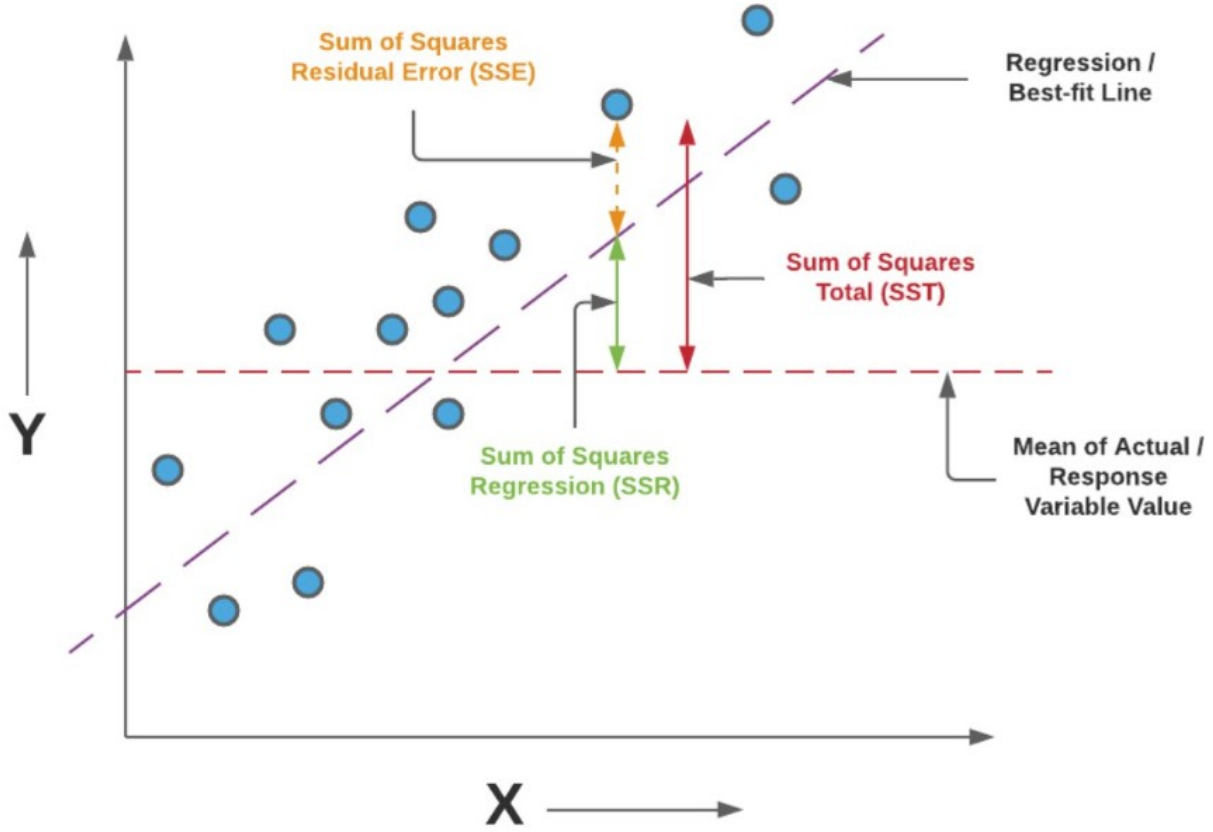
On linear regression construction and explanation. MatLab: fitlm



we will take a look at the **concepts** and **formula** of **f-test** and related **f-statistics** in linear regression models and understand **how to perform f-test and interpret f-statistics** in linear regression with the help of **examples**.

F-test and related F-statistics interpretation is key if you want to assess if the linear regression model results in a statistically significant fit to the data overall. An **insignificant F-test** determined by the **f-statistics value vis-a-vis critical region** implies that **the predictors have no linear relationship with the target variable**.

$H_o : \beta_1 = \cdots = \beta_2 = \beta_p = 0$

p - value is equal to the number of regression parameters

f = SSR / SSE (Mean sum of squares regression / Mean sum of squares error)

$f = (SSR / df_{SSR}) / (SSE / df_{SSE})$;

df_{SSR} = p - value is equal to the number of regression parameters

df_{SSE} =n-p-1 - the value is equal to the total number of records (n) minus the number of model's coefficients (p), so criterion statistics **$f = (SSR / p) / (SSE / n-p-1) \in F_{p,n-p-1}$**

$C_{crit}=F^{-1}_{p,n-p-1}(0.05)$;

If $f > C_{crit}$ **then** H_0 is rejected so $\exists\ \beta_i \neq 0$!

A larger F-statistic might indicate that the linear regression model accounts for a substantial portion of the total variance, while a smaller F-statistic suggests that the model might not explain much of the variance and thus, may not be seen as useful model.

Conclusion:

the t-test evaluates the significance of individual coefficients

F-test assesses the overall model significance, testing whether at least one predictor is effectively non-zero.

Our actions: eliminate the "bad" variables in the model

vis-à-vis [vɪs ɑ vɪs] - по отношению к