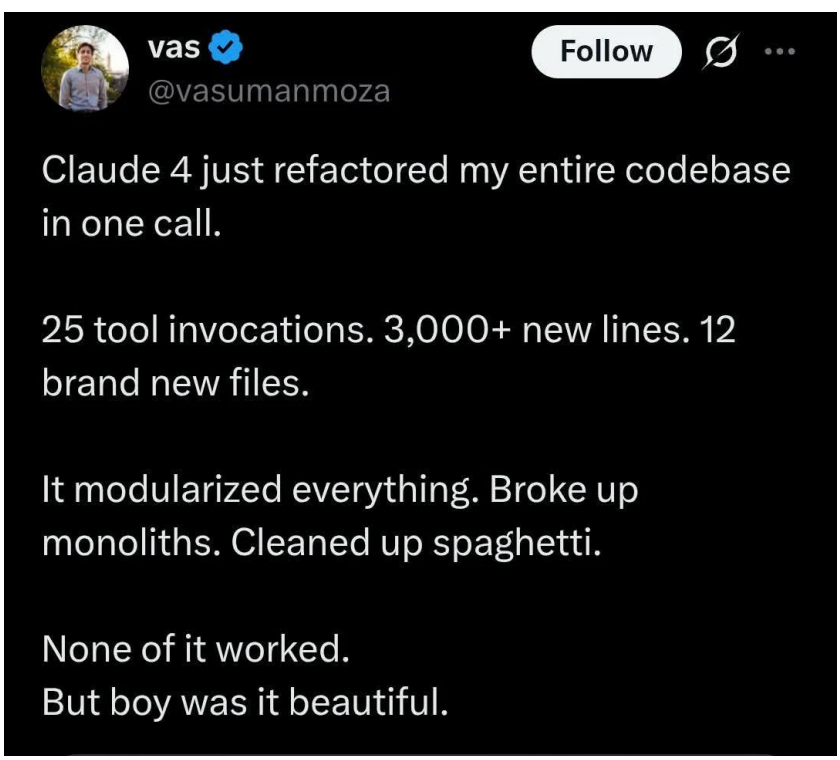
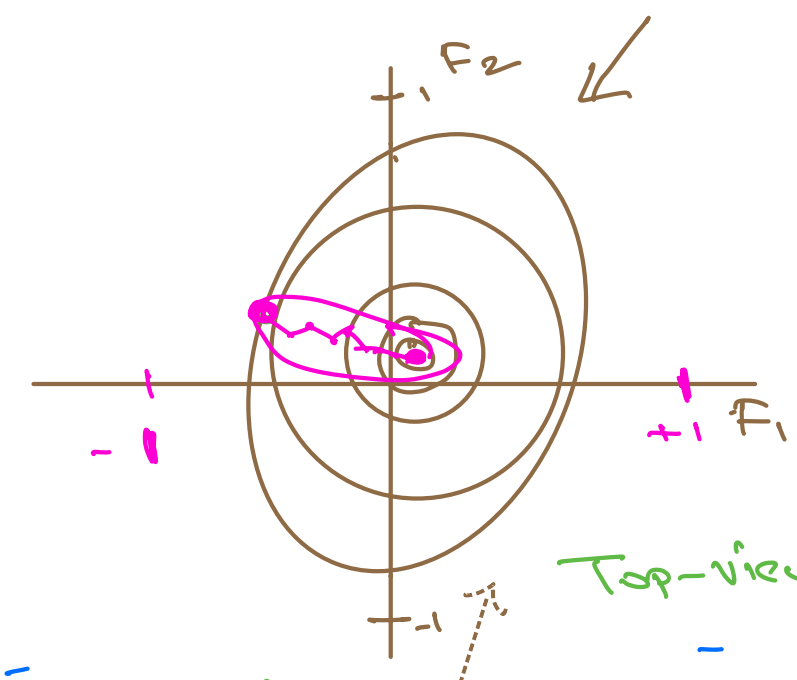
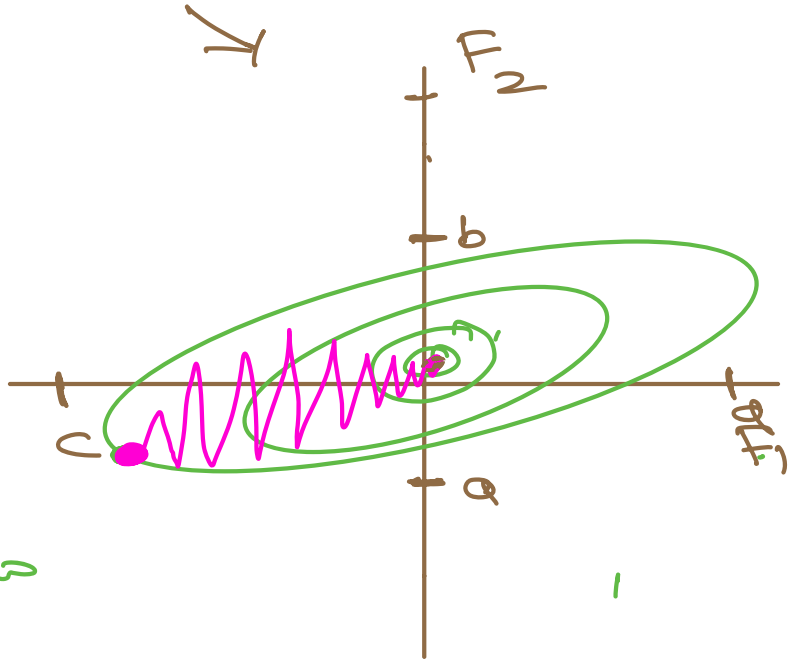
Agenda

- Feature Scaling impact on Loss Geometry
- Adjusted R2-score
- Sklearn vs statsmodel
- Assumptions of Linear Regression
  - Linear Relationship ✓
  - No multi-collinearity ✓
  - Normal Distribution of Residuals
  - Homoscedasticity
  - No auto-correlation

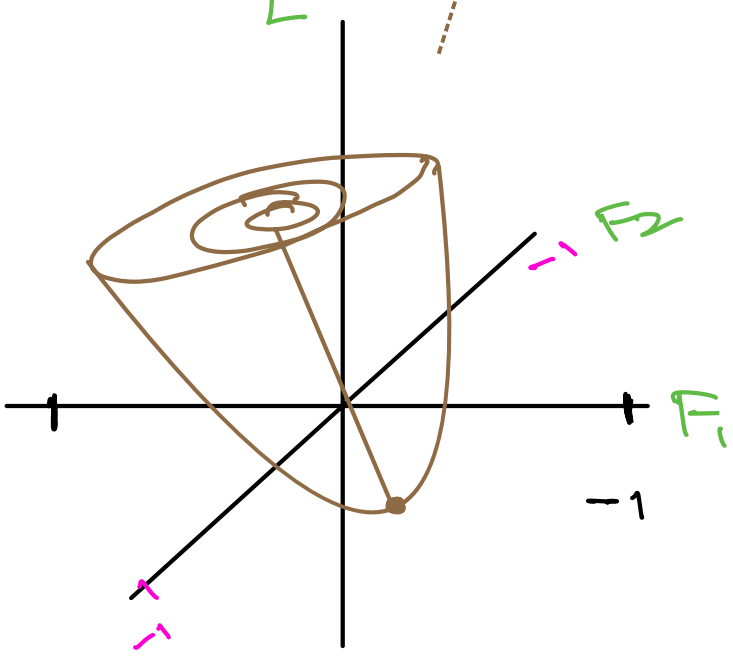


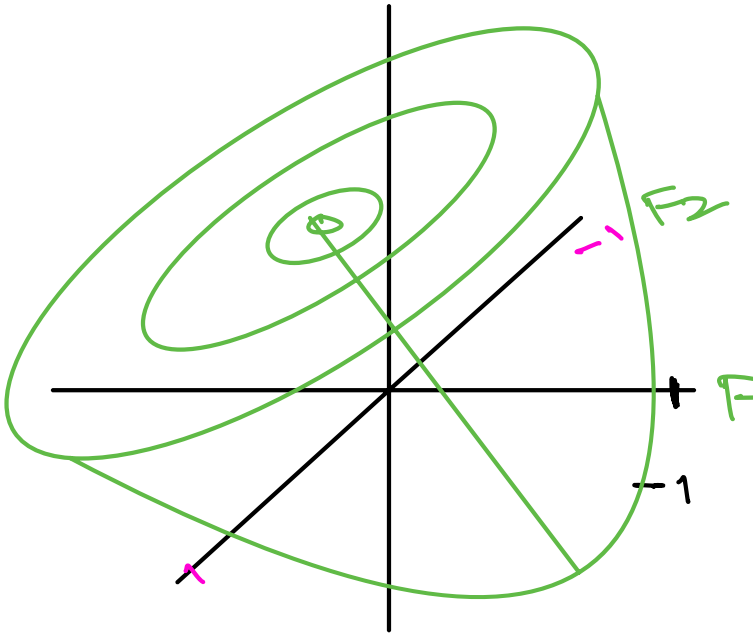
Feature Scaling impact on Loss Fx Geometry





Top-view





3d View of Loss

Adjusted R2- score

Case 1	Case 2
R2- score = 0.90	R2- score = 0.91
num-features = 2	num-features = 10

Occam's Razor

- When you have comparable performance, always pick the simple Version

Note: R2 score either 'increase or stay same when features are 'increased

'If we add 'irrelevant feature R2- score stays same

Adjusted R2_score

$$1 - \left[\frac{(1 - R\_2)(m - 1)}{(m - d - 1)} \right]$$

num_row

num_feature

$(-\infty, 1)$

$(0, 1)$

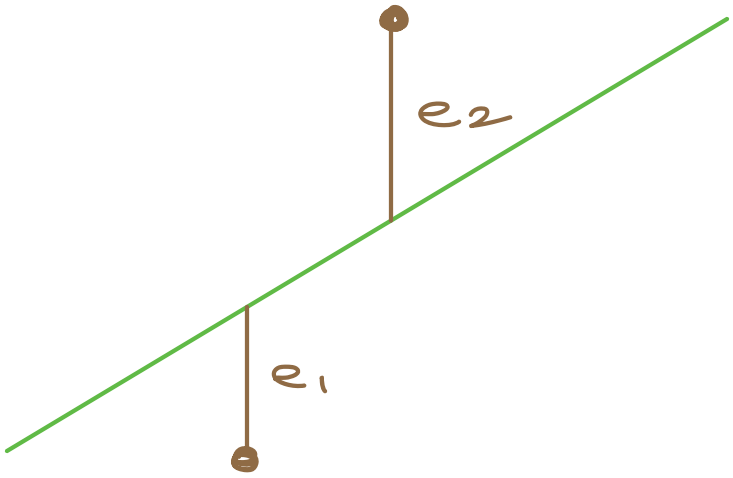
$d + 1$

If this dimension does not add significant value in Numerator adjusted R2_score will Reduce

if we add irrelevant feature adj r2_score will decrease

(H.W): Implement adj R2_score in our LinearRegression Class

Statsmodel



Ordinary Least Squares

$$\min \left(\text{Sum} \quad y \quad \text{Errors}^2 \right)$$

Stats Model	Sklearn
○ CI for features	○ Less concerned with statistical detail
○ F-statistic (Model)	
○ P-Value (individual feature)	○ Feature-scaling
○ Residual Plot	○ Regularization
○ QQ Plot	○ Cross-Validation
	○ Eval-metric

Assumptions of Linear Regression

1. Linear Relationship
2. No multi-collinearity
3. Normal Distribution of Residuals
4. Homoscedasticity
5. No auto-correlation

Pre-Read