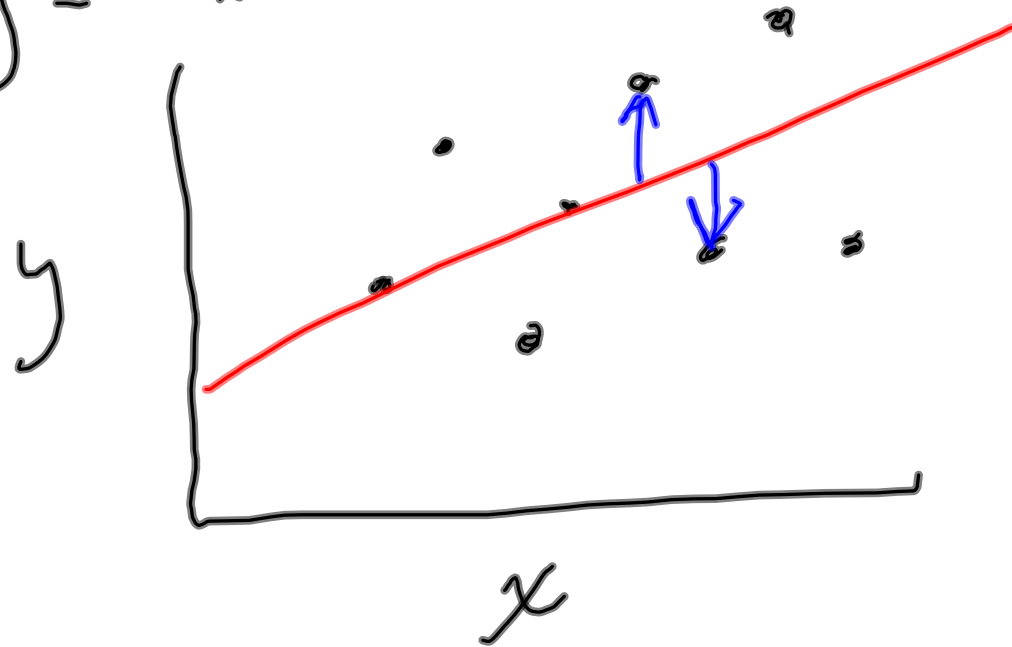
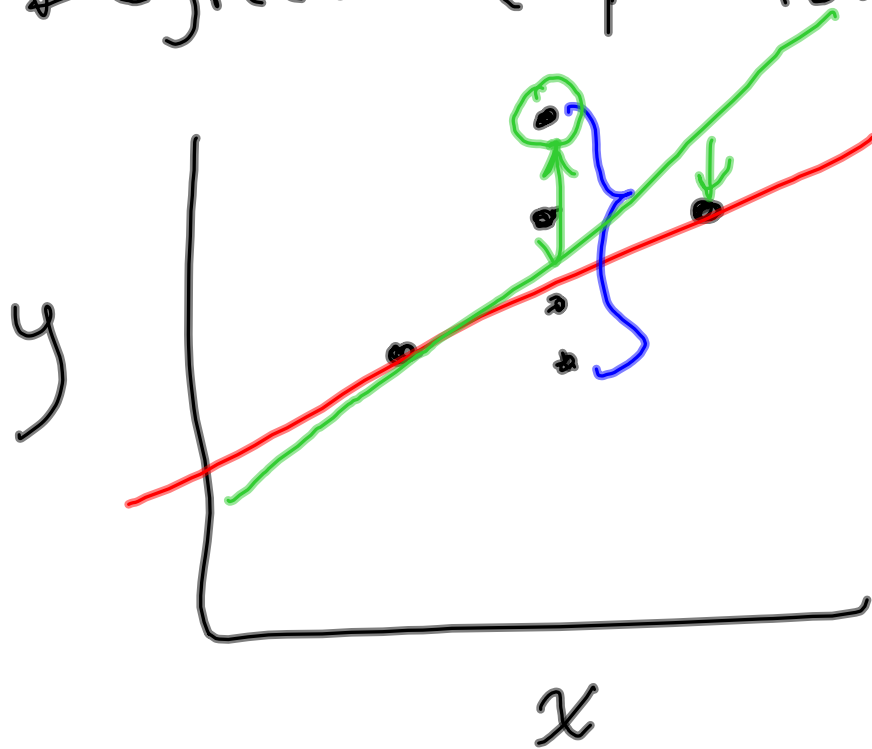


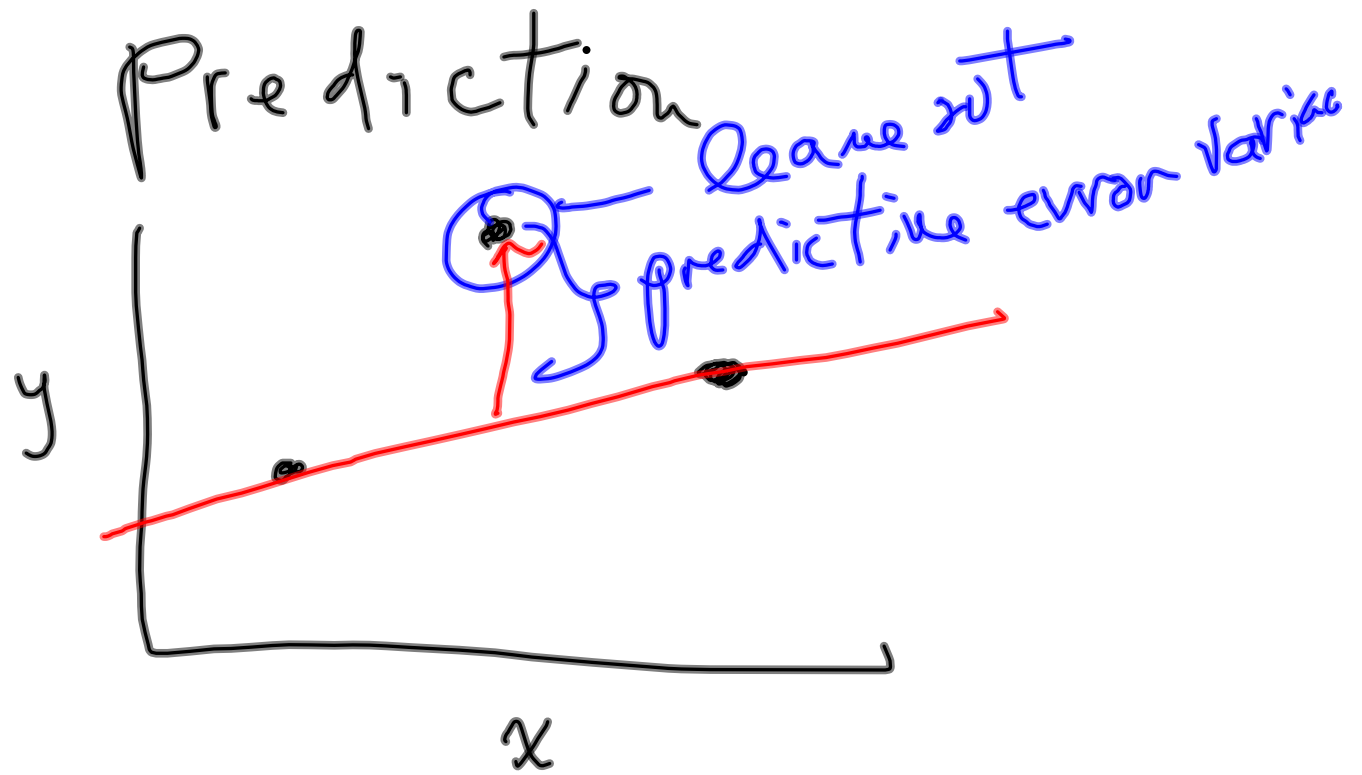
Regression

$$y = ax + b + \epsilon$$

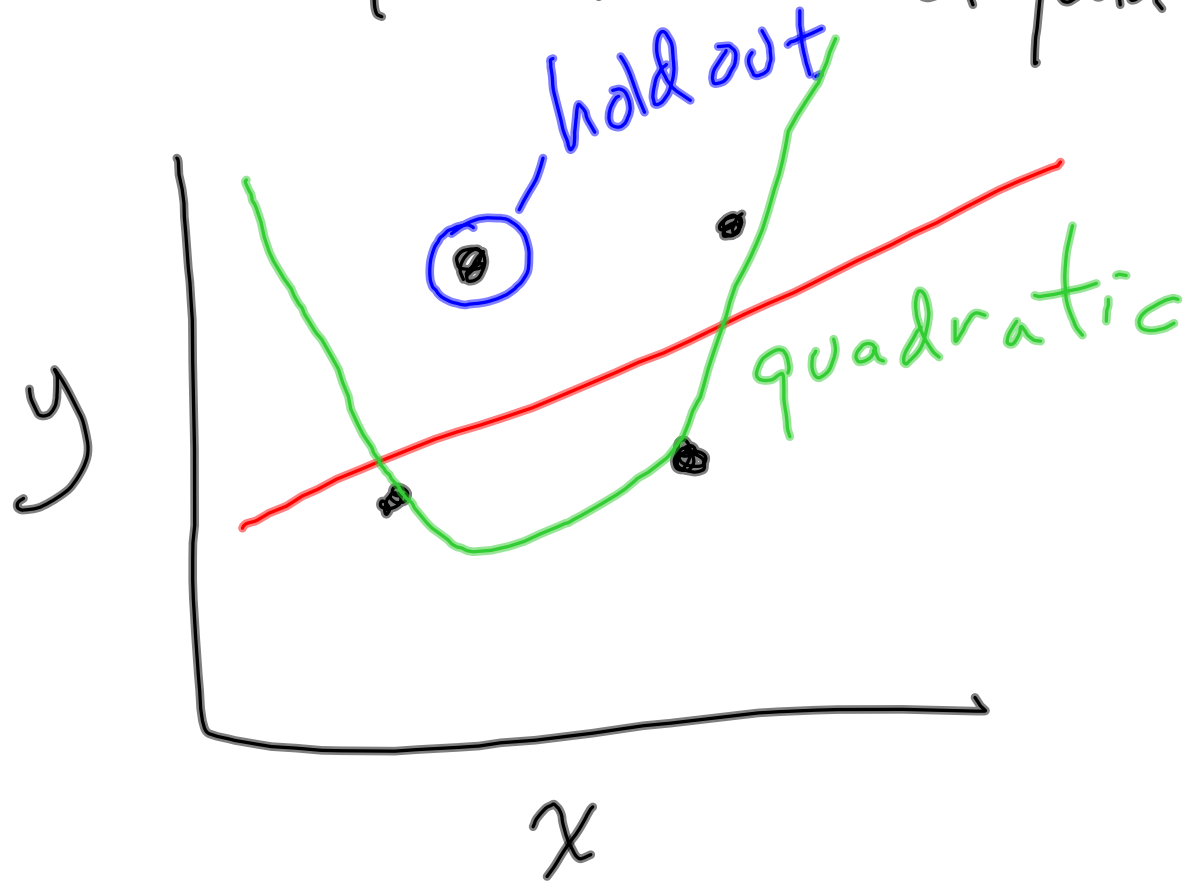


# Degrees of Freedom





# Choice between linear and quadratic

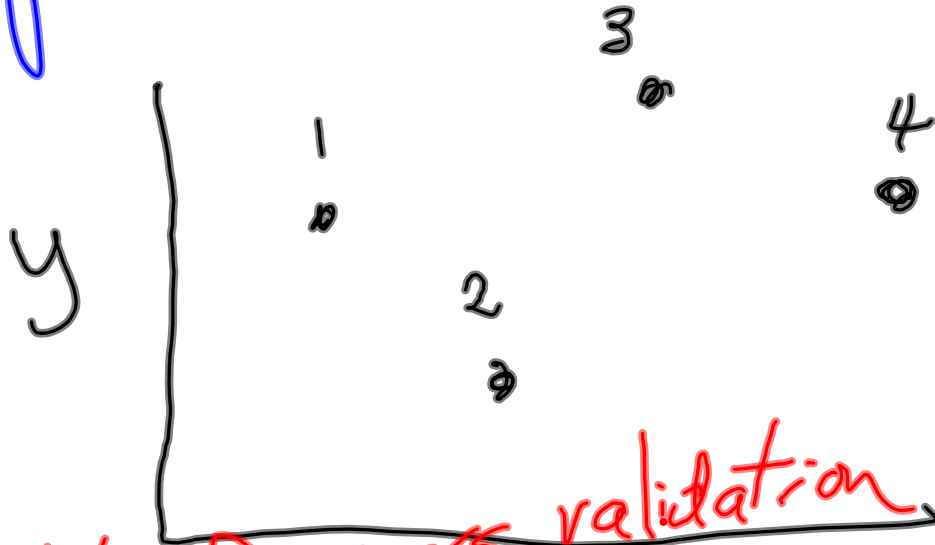


How do we select the data  
point to leave out?

If we only leave out one  
data point, isn't that  
a small sample?

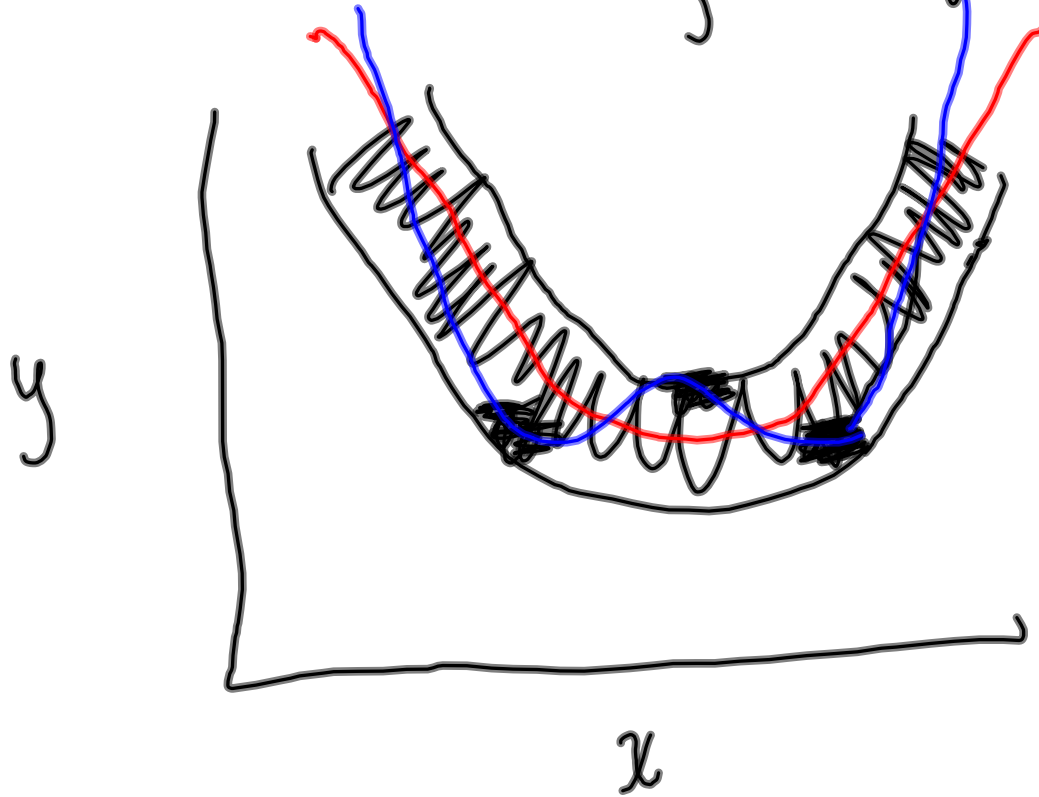
"Jack knife"

leave out each data point in turn.



jackknife cross validation  
prediction error  $\chi$  is  $\sum \epsilon_p^2$

Overwhelming sample size



Table

order of polynomial

	$r_2$	jackknife $r_2$
1	.4	.35
2	.5	.40
3	.6	.45
4	.7	.40
...		