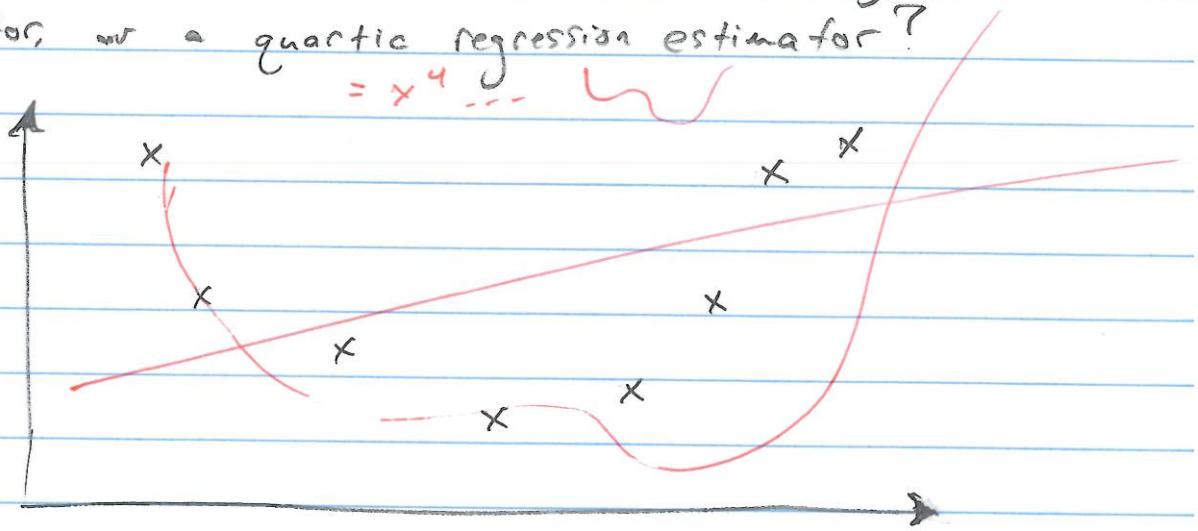
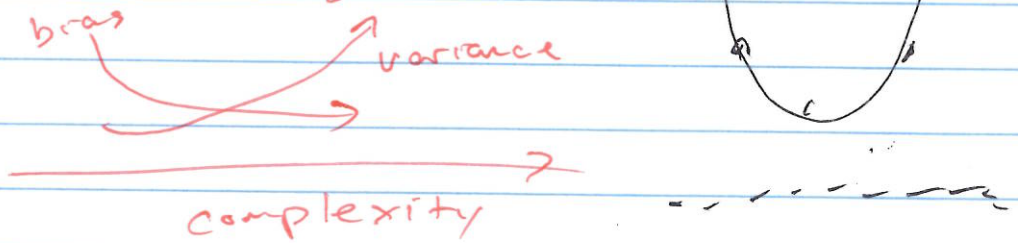
bias = $|x - E(x)|$

Given X generated by a noisy quadratic function, which is more biased, a linear regression estimator, or a quartic regression estimator?

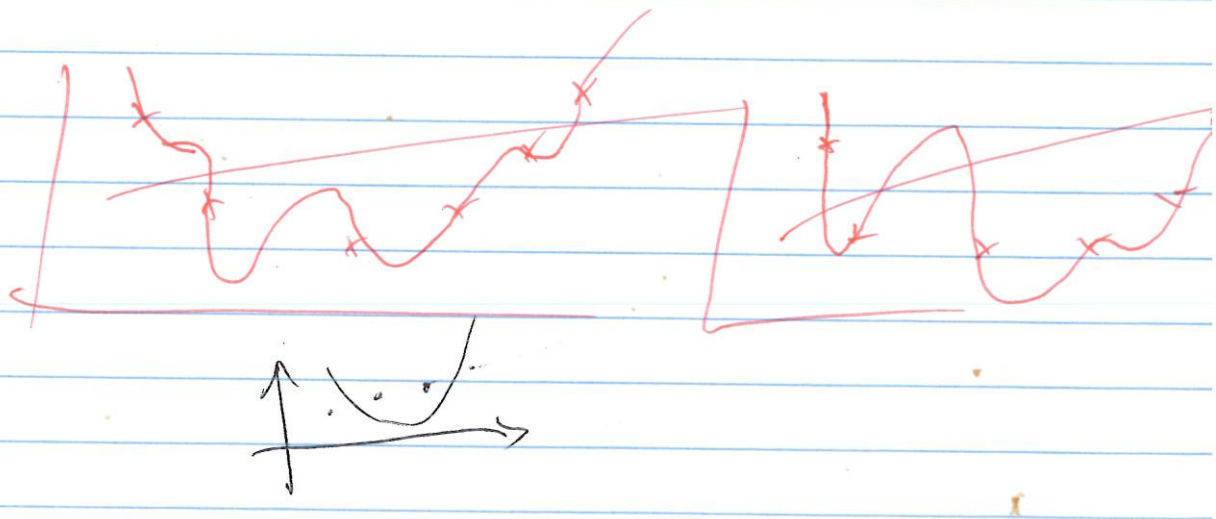


more bias — Linear regression



Which will have more variance?

Quartic



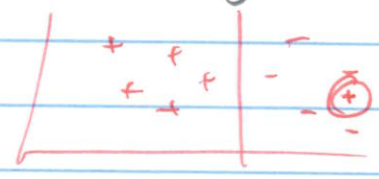
How do we control model complexity with...

neural nets?

change size of hidden layer

more units → more complexity

K nearest neighbors?



K ↓ complexity ↑

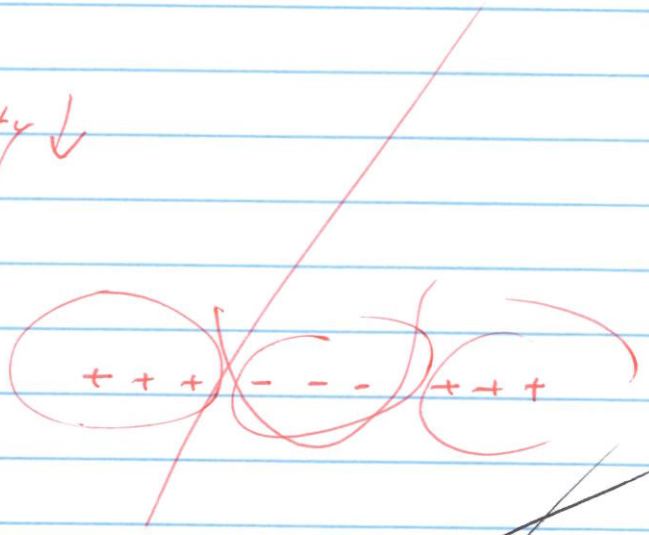
Logistic regression?

Regularization

↗ ↑ complexity ↓

SVM?

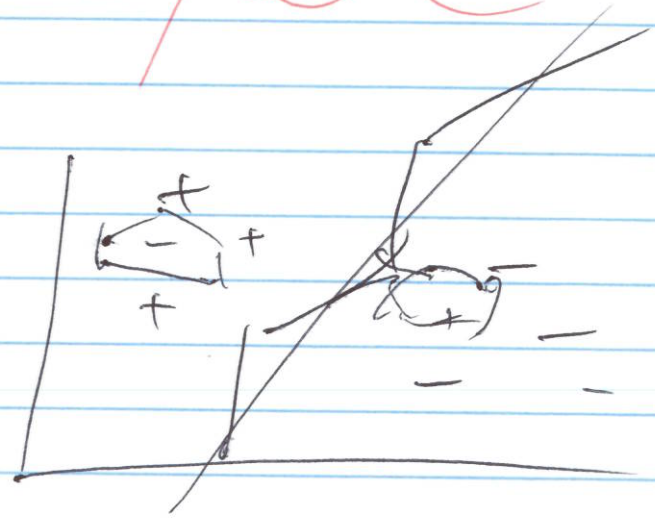
Kernel functions



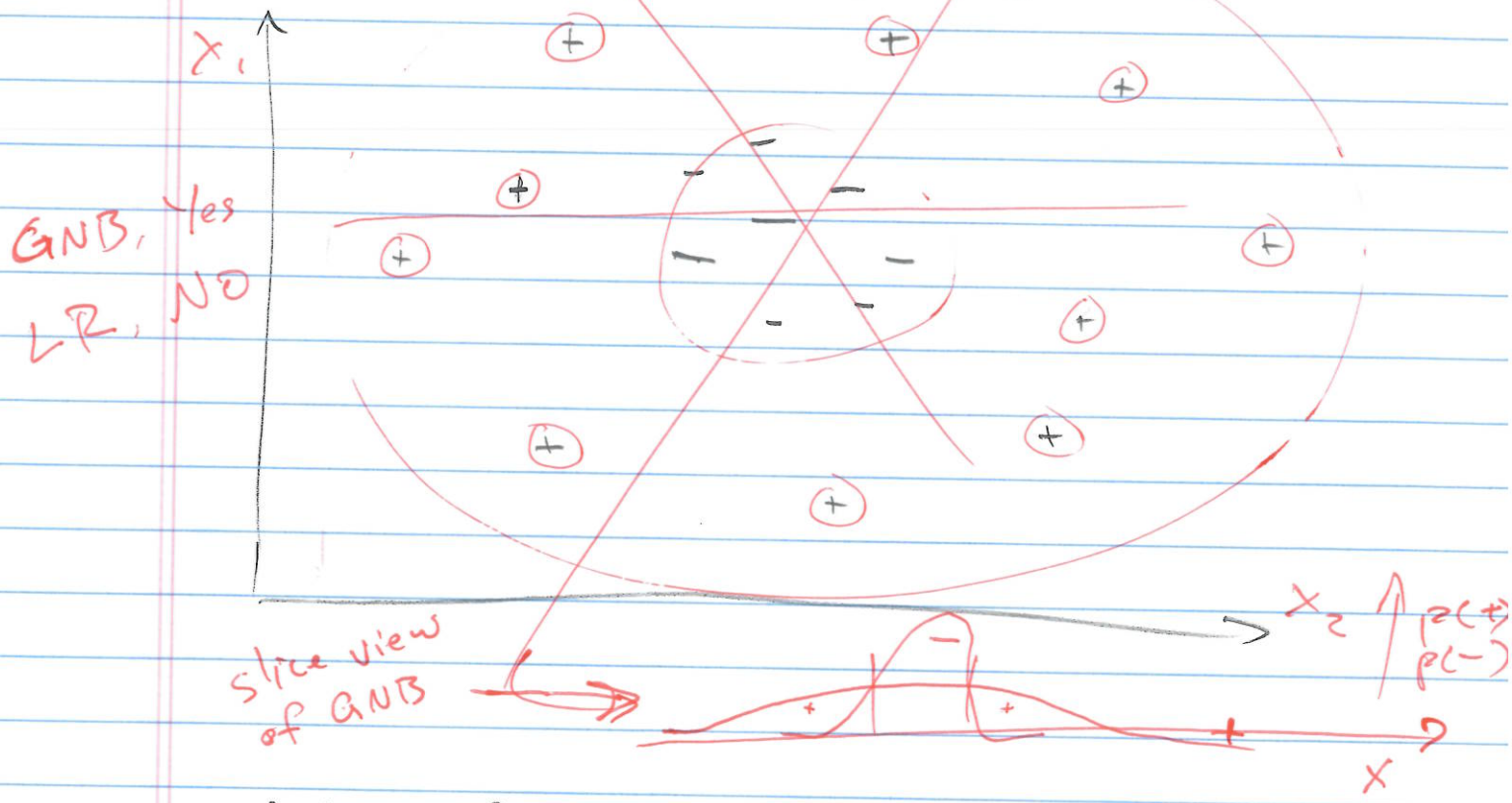
Decision Trees?

prune

or stop early



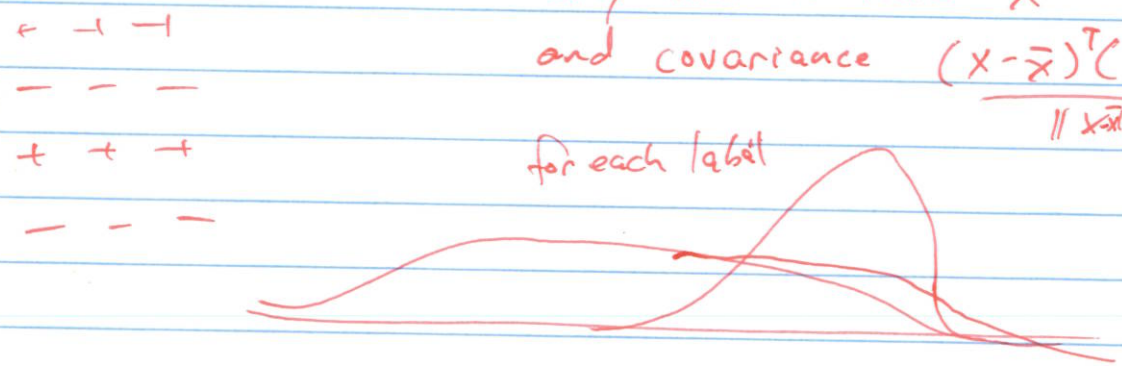
Can Gaussian Naive Bayes classify this data set? How about logistic regression?



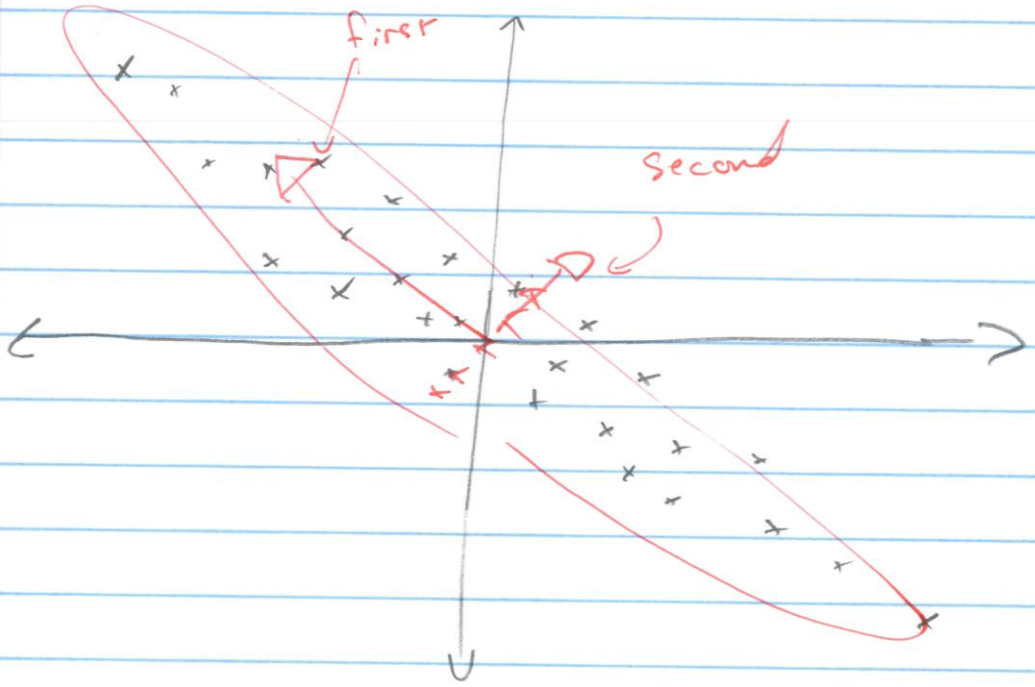
Which is faster to train, GNB or logistic regression?

GNB

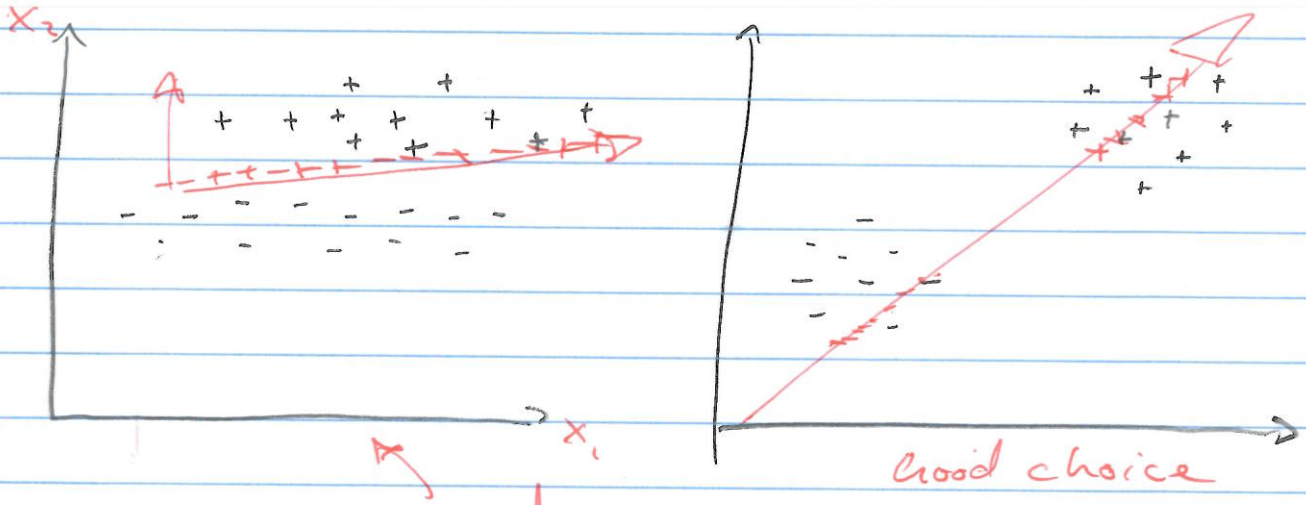
→ only need to calc $\bar{X}$ and covariance $\frac{(X-\bar{X})^T(X-\bar{X})}{\|X-\bar{X}\|}$ for each label



Approximately what are the first two principal components of this data set?



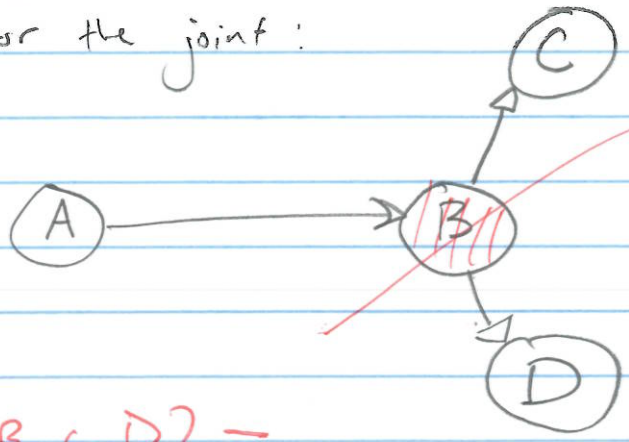
We want to reduce the dimensionality of a dataset for a classification task. In which case is PCA a good choice?



$$\max \frac{w^T X^T X w}{|w^T w|}$$

Everybody's favorite!

Factor the joint:



$P(A, B, C, D) =$

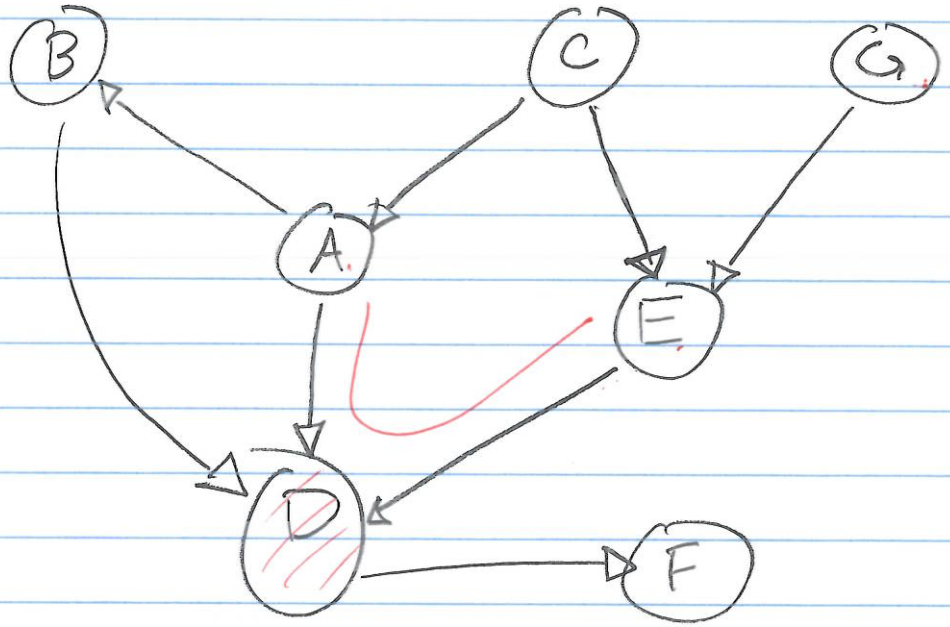
$P(A)P(B|A)P(C|B)P(D|B)$ ✓

I tell you $P(A=T \mid B=F) = 0.2$

Can you calculate $P(A=F \mid B=F, D=T)$?

$= 0.8$

Because $A \perp D \mid B$



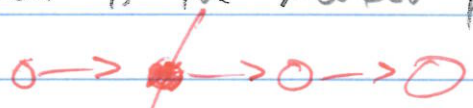
Markov blanket for C? F? B?

AEG D ADE

A ⊥ G? ✓ A ⊥ G | C? ✓ A ⊥ G | D? X

HMMs!

What is the Markov property of stochastic processes?



Future only depends on the present, not past sequence

During EM, what do we calculate in the E-step?

X = obs, Y = states

$$\alpha = P(X_1 \dots X_t, Y_t)$$

$$\beta = P(X_{t+1} \dots X_T, Y_{t+1})$$

M-step?

transition parameters,

emission,

initial

We have data X, initial state distn. π, transition matrix A, emission matrix B,

α where $\alpha_t = P(X_1 = x_1, \dots X_t = x_t, Y_t \mid A, B, \pi)$

Derive $P(X_t \mid X_1 = x_1, \dots X_{t-1} = x_{t-1})$

Exercise for the reader