3. Introduction to Non-linear Classification

Introduction to Non-linear Classification

0:00 / 0:00 1.50x

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Counting Dimensions of Feature Vectors

1/1 point (graded)

Let $x \in \mathbf{R}^{150}$, i.e. $x = \left[x_1, x_2, \ldots, x_{150}\right]^T$ where x_i is the i-th component of x. Let $\phi\left(x\right)$ be an **order** 3 polynomial feature vector. This means, for example, $\phi\left(x\right)$ can be

$$\phi\left(x\right) = \big[\underbrace{x_1, \ldots, x_i, \ldots, x_{150}}_{\text{deg 1}}, \underbrace{x_1^2, x_1x_2, \ldots, x_ix_j, \ldots x_{150}{}^2}_{\text{deg 2}}, \underbrace{x_1{}^3, x_1{}^2x_2, \ldots, x_ix_jx_k, \ldots, x_{150}{}^3}_{\text{deg 3}}\big] \qquad \text{where } 1 \leq i \leq j \leq k \leq 150.$$

Note that the components of $\phi\left(x\right)$ forms a basis of the space of all polynomials with zero constant term and of degree at most 3.

What is the dimension of the space that $\phi\left(x\right)$ lives in? That is, $\phi\left(x\right) \in \mathbb{R}^d$ for what d?

Hint: The number of ways to select a multiset of k non-unique items from n total is $\binom{n+k-1}{k}$. For example, if a ball can be any of 3 colors, then the number of color configurations of 2 balls is $\binom{3+2-1}{2} = \binom{4}{2} = 6.$

$d =$ 585275 Answer: 585275

Solution:

For each of the feature transformations (power 1, power 2, power 3), there are n-multichoose-power combinations. Thus $\binom{150}{1} + \binom{151}{2} + \binom{152}{3} = 585275$. **Remark:** We see that the dimension of the space that the feature vectors live grows quickly as a function of d, the dimension we started with if $x \in \mathbb{R}^d$.

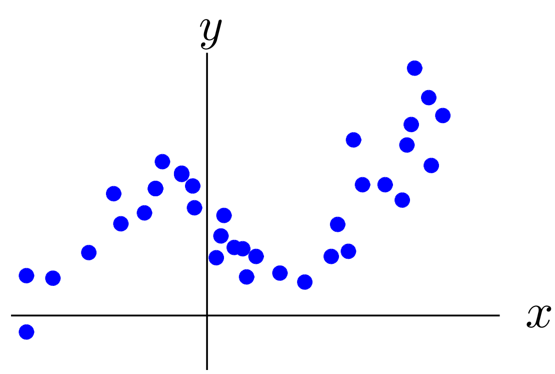
Submit You have used 1 of 3 attempts

Answers are displayed within the problem

Regression using Higher Order Polynomial feature

1/1 point (graded)

Assume we have n data points in the training set $\{(x^{(t)}, y^{(t)})\}_{t=1,\ldots,n}$ where $(x^{(t)}, y^{(t)})$ is the t-th training example:



We want to find a non-linear regression function f that predicts y from x, given by

$$f(x; \theta, \theta_0) = \theta \cdot \phi(x) + \theta_0$$

where $\phi(x)$ is a polynomial feature vector of some order. What (loosely) is the minimum order of $\phi(x)$?

3 Answer: 3

Solution:

The relationship between y and x can be roughly described by a cubic function, so a feature vector $\phi(x)$ of minimum order 3 can minimize structural errors.

Submit You have used 1 of 2 attempts

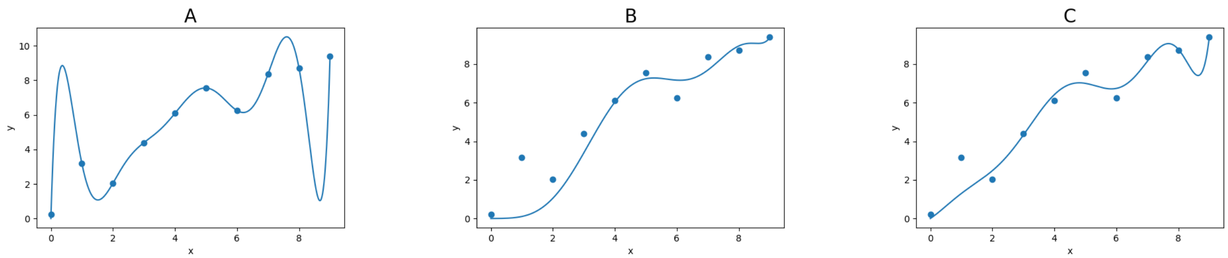
Answers are displayed within the problem

Effect of Regularization on Higher Order Regression

2/2 points (graded)

Let us go back to explore the effect of regularizaion on Higher Order regression.

The three figures below show the fitting result of a 9th order polynomial regression with different regularization parameter λ on the same training data.



Which figure above corresponds to the smallest regularization parameter λ?

- (•) A
- ( ) B
- ( ) C

✔

Which figure corresponds to the largest regularization parameter λ?

- ( ) A
- (•) B
- ( ) C

✔

Solution:

The effect of regularization is to restrict the parameters of a model to freely take on large values. This will make the model function smoother, leveling the 'hills' and filling the 'vallyes'. It will also make the model more stable, as a small perturbation on x will not change y significantly with smaller $\|\theta\|$.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2 weeks):Lecture 6. Nonlinear Classification / 3. Introduction to Non-linear Classification

Add a Post

Show all posts by recent activity

- Enquiry per Question 1 — 2
  Why the calculation has to be done by the (n,k) formula? If I consider it as a counting problem divided by sequence (as sequence does not matter for polynomial) , why isn't th...
- Non-linear regression function - minimum order Polynomial feature — 4
  I am confuse with something that looks obvious. Should an order 2 Polynomial feature be not linear? I would say so since we introduce non-linear terms (such as x1x2 or x1^2...
- Q1 Counting Dimensions of Feature Vectors, another way to look at it — 4

Segment 6.3 notes

Community TA

3

Confused in regard to the regularization parameter

2

Counting Dimensions of feature vector

https://en.wikipedia.org/wiki/Multinomial_theorem This context suits better than the example give below

1

Resource to understand effect of λ

Watch this to understand the last question: https://youtu.be/Q81RR3yKn30?t=452

1

Counting Dimensions of Feature Vectors: Hint?

I solved this easy one on my third attempt, only because I was trying to fit somehow the Hint in the solution process. After two attempts, I just thought about the problem and...

3

Transcription error

2:29: phi of x instead of 5x

Community TA

1

Q1 Counting Dimensions, fail to understand the hint

Could anyone give a hint on how to understand the given hint? I used to apply other way to compute dimension of feature vectors, which seemingly is wrong here. (Maybe so...

3

Video

Would be nice to see the actual resulting accuracy of each polynomial regression using the leave one out method

1

Learn About Verified Certificates