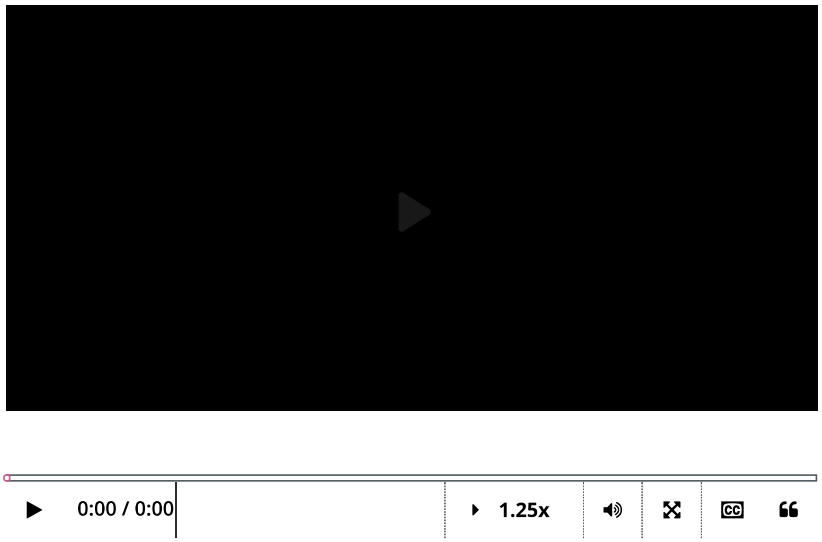
4. Motivation for Kernels: Computational Efficiency

Exercises due Oct 14, 2020 05:29 IST *Completed*

Motivation for Kernels: Computational Efficiency



Start of transcript. Skip to the end.

The idea is that you can take inner products

between high dimensional feature vectors

and evaluate that inner product very cheaply.

And then, we can turn our algorithms

into operating only in terms of these inner products.

Video
Download video file

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

Kernels as Dot Products 1

1/1 point (graded)

Let us go through the computation in the video above. Assume we map x and $x' \in \mathbb{R}^2$ to feature vectors $\phi(x)$ and $\phi(x')$ given by

$$\begin{aligned}\phi(x) &= \left[x_1, x_2, {x_1}^2, \sqrt{2}x_1x_2, {x_2}^2\right] \\ \phi(x') &= \left[x'_1, x'_2, {x'_1}^2, \sqrt{2}x'_1x'_2, {x'_2}^2\right].\end{aligned}$$

Which of the following equals the dot product $\phi(x) \cdot \phi(x')$?

- $x \cdot x'$
- $x \cdot x' + (x \cdot x')^2$ ✔
- $(x \cdot x')^2$
- $2(x \cdot x')^2$
- None of the above

✔

Solution:

Expand $\phi(x) \cdot \phi(x')$ to get

$$\phi(x) \cdot \phi(x') = x_1x'_1 + x_2x'_2 + {x_1}^2{x'_1}^2 + 2x_1x'_1x_2x'_2 + {x_2}^2{x'_2}^2$$

$$= (x_1 x_1' + x_2 x_2') + (x_1 x_1' + x_2 x_2')^2$$
$$= x \cdot x' + (x \cdot x')^2.$$

Remark: Notice the coefficient $\sqrt{2}$ of the $x_1 x_2$ terms is necessary for rewriting $\phi(x) \cdot \phi(x')$ as the function above of $x \cdot x'$.

Submit You have used 1 of 1 attempt

Answers are displayed within the problem

Kernels as Dot Products 2

1/1 point (graded)
Which of the following feature vectors $\phi(x)$ produces the kernel

$$K(x, x') = \phi(x) \cdot \phi(x') = x_1 x_1' + x_2 x_2' + x_3 x_3' + x_2 x_3' + x_3 x_2'$$

- $\phi(x) = [x_1, x_2, x_3]$
- $\phi(x) = [x_1 + x_2 + x_3]$
- **(selected)** $\phi(x) = [x_1, x_2 + x_3]$
- $\phi(x) = [x_1 + x_3, x_1 + x_2]$

✔

Solution:

Directly expand to see the answer. The fact that there are mixed terms in the kernel, e.g. $x_2 x_3'$, indicates that some coordinates of the feature vector must be mixed, i.e. contain different x_i's.

Submit You have used 1 of 1 attempt

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2 weeks):Lecture 6. Nonlinear Classification / 4. Motivation for Kernels: Computational Efficiency

Add a Post

Show all posts by recent activity

? why kernel is called kernel? — 10
Hi, In kernel function, what is the meaning of kernel? Why it is called kernel?

? What does x' mean? — 2
According to the last question, if we know phi x then we know about phi x'?

Why Kernel? — 1
Why do we calculate kernel with two feature vectors? where is this applied?

? Potential issue with the question "Kernels as Dot Products 1" — 2

Shouldn't the vectors in question "Kernels as Dot Products 1" be transposed in order for the answer you point as correct, to be correct? I... 2

? technical problem: mistakenly clicked on "save" after submitting a correct answer 2
After submitting a correct answer using my one available attempt, I mistakenly hit the "save" button to see what it would do. It seems to...

< Previous Next >

/