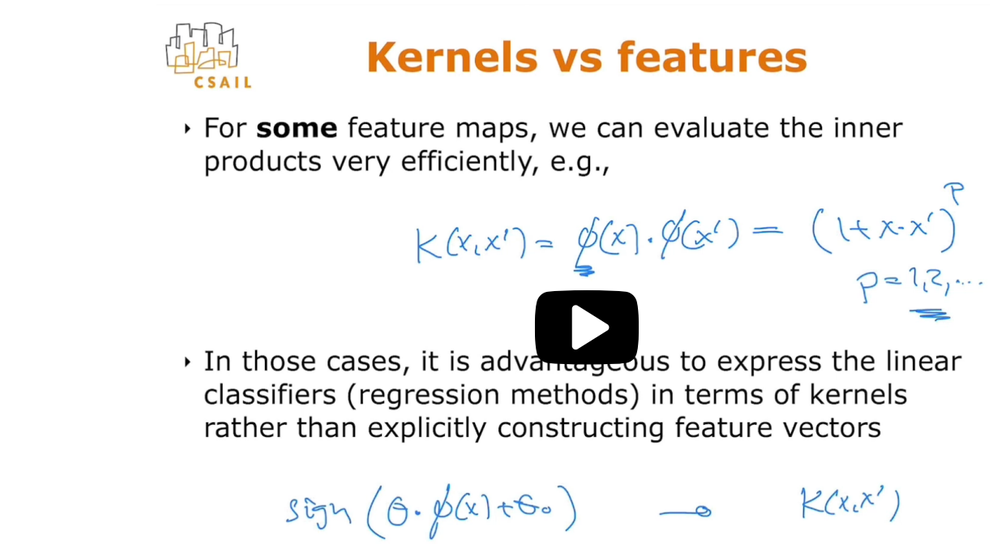
4. Motivation for Kernels: Computational Efficiency

Motivation for Kernels: Computational Efficiency



takes theta dot f of x five plus theta naught

and tries to classify that-- that's how a linear classifier.

Somehow, we wish to turn that inter-classifier

that only depends on those inner products

operates in terms of kernels.

And we'll do that in the context of kernel perception

just for simplicity.

But it applies to any linear method

that you've already learned.

End of transcript. Skip to the start.

▶ 5:34 / 5:34 Speed 1.50x

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

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

Remark: Notice the coefficient $\sqrt{2}$ of the x_1x_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

$$K(x, x') = \phi(x) \cdot \phi(x') = x_1 x'_1 + x_2 x'_2 + x_3 x'_3 + x_2 x'_3 + x_3 x'_2$$

(Choose all that apply.)

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

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 ▼

?	Kernel method for regression or classification? Is the Kernel method for feature transformation used only in classification problems or also for regression?	3
☑	What does the dot indicate in the first question in the options ? is it dot product ? x.x'	3
?	[STAFF] Who is x' ??? Can you tell me who is x' ?	3
💬	[staff] Little change for 2nd question. The answer options are radio buttons, so only one can be selected. The line "Choose all that apply." looks leftover. Community TA	4
💬	Tiny terminology nitpick Community TA	1
?	X1 + X3 When you create x1+x3 as new feature, for example, if salary and rental income are x1 and x3 if you add x1+x3, then you will get total income, if you multiply x1 with (x1+x3) , ...	4
?	Computational cost of additional features Is not just in creating them by inner product, but also in calulating parameter updates, when there are so many features correct ? Inaddition, how a dot product of (x1, x2, x3) ...	2

Learn About Verified Certificates