



Lecture 7 | Kernels | Stanford CS229

🕒 Date de création	@January 9, 2025 1:33 PM
🏷️ Étiquettes	

Questions

- What are the key ideas?
- What terms or ideas are new to me?
- How would I define them?
- How do the ideas relate to what I already know?
- Should not assume that the profesor are always correct. Asking appropriate questions.
- What are the good ideas?
- Do the ideias have other applications?

▼ What means $||w||$?

$$||x|| = \sqrt{x_1^2 + x_2^2}$$

Generative models learn backwards (features to class).

▼ How to make Kernels

Lecture

Outline

- Optimization problem
- Representer Theorem
- Kernels
- Examples Of Kernels

One simple reason why SVM is great is that, it requires little use of parameters updates, you run and works.

SVMs is a combination of Margin Classifier and Kernels, and all math in SVMs seems contribute to increase the difference between a training point and the geometric margin.

When working with SVMs the number of dimensions could be a problem, in order to deal with, we make w a linear combination of the training examples: $w = \sum_{n=1}^m \alpha_n x^{(n)}$

In order to use Kernel Trick we need assure something, write your algorithm in terms of a **inner product**. Represente your input features in a high dimensionl and then compute $K(x, z) = \Phi(x)^T \Phi(z)$. And Finnaly replace (x, z) for $K(x, z)$.

Also, all this math is to simplyfy the couts, use less dimensations, reduce O complexity.

L1_norm make the margin more robust to outliers.

Vocab

- Attributes: we will call the “original” input value the input attributes of a problem
- Features: when the original input is mapped to some new set of quantities $\phi(x)$

- Gaussian Kernel: A ***Gaussian Kernel*** refers to a mathematical function used to model local deformation in computer science.
- Mercer's Theorem: In mathematics, specifically functional analysis, **Mercer's theorem** is a representation of a symmetric positive-definite function on a square as a sum of a convergent sequence of product functions.
- Representer Theorem: In computer science, this representation theorem is one of several results that show the minimizer of a fixed empirical risk value \times **(Finite Linear)** definition of a reproducing kernel Hilbert space.
- Dual Optimization Problem: The duality principle says that the optimization can be viewed from 2 different perspectives.
- Deductive Bias: the set of assumptions that the learner uses to predict outputs of given inputs that it has not encountered.
- Knuth-Morris-Pratt Algorithm:

More

- <https://www.geeksforgeeks.org/representer-theorem/>
- <https://pages.stat.wisc.edu/~mchung/teaching/MIA/reading/diffusion.gaussian.kernel.pdf.pdf>
- https://en.wikipedia.org/wiki/Mercer%27s_theorem
- <https://medium.com/@sathvikchiramana/svm-dual-formulation-7535caa84f17>
- [The Kernel Trick - THE MATH YOU SHOULD KNOW!](#)
- <https://www.youtube.com/watch?v=Q7vT0--5VII>