

During the course of the last week, we have discussed about inner products in their respective vector spaces and how they are uniquely defined. A norm induced by the inner product will satisfy extra properties as compared to the other norms. A Grammian operator used in finding the point on a subspace closest to a given point employs the use of inner products. This is technically called linear approximation in a Hilbert space. This concept explores an easy method to calculate the closest distance using the properties of inner products. Using this algorithm, we have found out the structure of a polynomial closest to the exponential function. We have also seen examples (the last example in the lecture 5 notes) showing how weights can be assigned as part of the algorithm (to reward/punish) during each iteration for finding the best fitting polynomial. This incentive-driven method will make sure that the machine learns faster and with greater accuracy. A set of bases need not be orthogonal to each other, but if they are (orthogonal bases), then these help save computational time in the algorithms of linear approximation by reducing the Grammian to an identity matrix. The orthogonal bases make calculations a lot easier.

The combined concepts of orthogonal bases, inner product and linear approximation will pave the way for a better intuitive understanding of machine learning. The exact use of linear algebra in machine learning is still very cloudy for me as we have not yet covered its applications. On googling, I found that if I can understand machine learning methods at the level of vectors and matrices, I will be able to improve my intuition for how and when they work in the case of machine learning. Matrices help us to look at all the data as a single entity. They also say that it has been observed through practice that representing large sets of data (in the form of vectors and matrices) help us visualize the data better. Linear algebra provides the computational engine for the majority of machine learning algorithms. Many of the problems in ML are about data fitting: Given a subspace  $S$  and a point  $p$  in a given scalar product space, find on  $S$  a point closest to  $p$ . Figuring out this 'best fit' is important for classifying unlabelled data points. With these in mind, I look forward to the next unit which deals with the concepts of linear estimation.