Abstract:

We develop a general duality between neural networks and compositional kernels. We show that initial representations generated by common random initializations are sufficiently rich to express all functions in the dual kernel space. Hence, though the training objective is hard to optimize in the worst case, the initial weights form a good starting point for optimization. Our dual view also reveals a pragmatic and aesthetic perspective of neural networks and underscores their expressive power. Joint work with Roy Frostig and Yoram Singer