データの正規化

データの正規化は、学習させる上で、重要なデータと重要でないデータとの差をはっきりさせる、または、値が大きくならないようにするという意味だけではない。

誤差関数の話の時に、神の分布との差を測る話をした。

データの正規化は、その神の分布との差の測り方を決める重要な役割を果たしている。

具体的に、‘Laplace‘の正規化を考えてみる。

‘Laplace‘の正規化では、平均が０、分散が１になるようにデータを正規化している。つまり、１次モーメント（平均）、２次モーメント（分散）は、神の分布とだいたい一緒であると仮定し（無視）、３次モーメント以降で神の分布との差を測っているのである。

※高次モーメントの差が無くなるほど、グラフの小さいところまで、非常に近い形をしているということである。

つまり、データの正規化の時点で、あまり正答率が高くなくて良い（神の分布と違っても良い）ところは無視し、重要な部分だけを、神の分布との差を測れるようにしているのである。