機器學習HW4筆記

1. Logistic regression

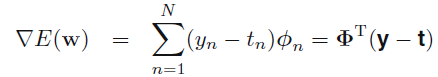
By Newton-Raphson update



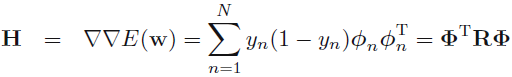
the posterior probability of class *C*1



the gradient of the error function with respect to **w**



the Hessian matrix



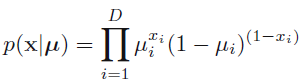
where the *N × N* diagonal matrix **R**



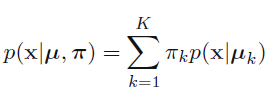
1. EM

**E step**

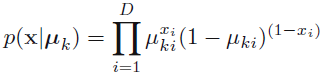
each pixel is governed by a Bernoulli distribution with parameter μi



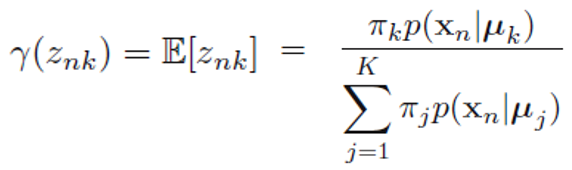
a finite mixture of these distributions



where

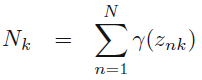


responsibility

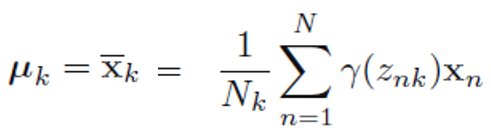


**M step**

the effective number of data points associated with component *k*



the mean of component *k* equal to a weighted mean of the data



the mixing coefficient for component *k* is given by the effective fraction of points in the data set explained by that component

