

5.40 For multi-class problems, the log-likelihood function is same as 4.108:

$$\ln p(\mathcal{D}|\mathbf{w}) = \sum_{n=1}^N \sum_{k=1}^K t_{nk} \ln y_n$$