# Chapter 4 - Linear Models for Classification

## 2. Probabilistic Generative Models

model class-conditional densities p(xICk) and class priors p(Ck), then use them to compute posterior probabilities p(CkIx) through Bayes' theorem

$$
\begin{aligned}
p(\mathcal{C}_k|\mathbf{x}) &= \frac{p(\mathbf{x}|\mathcal{C}_k)p(\mathcal{C}_k)}{\sum_j p(\mathbf{x}|\mathcal{C}_j)p(\mathcal{C}_j)} \\
&= \frac{\exp(a_k)}{\sum_j \exp(a_j)}
\end{aligned}
$$

$$a_k = \ln p(\mathbf{x}|\mathcal{C}_k)p(\mathcal{C}_k)$$

### 2.1. Continuous inputs

assumption: class-conditional densities - Gaussian

$$p(\mathbf{x}|\mathcal{C}_k) = \frac{1}{(2\pi)^{D/2}}\frac{1}{|\mathbf{\Sigma}|^{1/2}}\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\right\}$$

$$a_k(\mathbf{x}) = \mathbf{w}_k^{\mathrm{T}}\mathbf{x} + w_{k0}$$

$$
\begin{aligned}
\mathbf{w}_k &= \mathbf{\Sigma}^{-1}\boldsymbol{\mu}_k \\
w_{k0} &= -\frac{1}{2}\boldsymbol{\mu}_k^{\mathrm{T}}\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_k + \ln p(\mathcal{C}_k)
\end{aligned}
$$

### 2.2. Maximum likelihood solution

maximisation with respect to p(Ck) = Nk / N

maximisation with respect to µk = Σ(xi) / Nk

maximisation with respect to S = Σ(Ni * Si) / N