7.1 Linear Discriminant Analysis

The classification criterion for c_1 (or $\neg c_2$) can be expressed as follows:

$$
\begin{aligned}
p(c_1|x) &> p(c_2|x) \qquad |\text{apply Bayes' rule} && (1)\\
\frac{p(x|c_1)p(c_1)}{p(x)} &> \frac{p(x|c_2)p(c_2)}{p(x)} && (2)\\
\exp(-\frac{1}{2}(\vec{x}-\vec{\mu_1})^T\Sigma^{-1}(\vec{x}-\vec{\mu_1}))\, p(c_1) &> \exp(-\frac{1}{2}(\vec{x}-\vec{\mu_2})^T\Sigma^{-1}(\vec{x}-\vec{\mu_2}))\, p(c_2) \qquad |\log && (3)\\
-\frac{1}{2}(\vec{x}-\vec{\mu_1})^T\Sigma^{-1}(\vec{x}-\vec{\mu_1}) + \log p(c_1) &> -\frac{1}{2}(\vec{x}-\vec{\mu_2})^T\Sigma^{-1}(\vec{x}-\vec{\mu_2}) + \log p(c_2) && (4)\\
0 &< \vec{x}^T \underbrace{\Sigma^{-1}(\vec{\mu_1} - \vec{\mu_2})}_{w} \underbrace{-\frac{1}{2}(\vec{\mu_1}^T\Sigma^{-1}\vec{\mu_1} - \vec{\mu_2}^T\Sigma^{-1}\vec{\mu_2}) + \log\frac{p(c_1)}{p(c_2)}}_{b} && (5)
\end{aligned}
$$

The mathematical conversions lead to a linear equation in $\vec{x}$ where we can directly read the parameters w and b. In case of unequal covariances we cannot eliminate the term in front of the exponential equation in $p(x|c_k)$ which would lead us to a nonlinear result and therefore would require a nonlinear connectionist neuron. The shape of the decision boundary would bend in the direction of the data, whose class has the smaller variance.