LDA AS LINEAR CLASSIFIER

$$
\begin{aligned}
& \pi_k \cdot p(\mathbf{x}|y=k) \\
\propto \quad & \pi_k \exp\left(-\tfrac{1}{2}\mathbf{x}^T\Sigma^{-1}\mathbf{x} - \tfrac{1}{2}\boldsymbol{\mu}_k^T\Sigma^{-1}\boldsymbol{\mu}_k + \mathbf{x}^T\Sigma^{-1}\boldsymbol{\mu}_k\right) \\
= \quad & \exp\left(\log \pi_k - \tfrac{1}{2}\boldsymbol{\mu}_k^T\Sigma^{-1}\boldsymbol{\mu}_k + \mathbf{x}^T\Sigma^{-1}\boldsymbol{\mu}_k\right) \exp\left(-\tfrac{1}{2}\mathbf{x}^T\Sigma^{-1}\mathbf{x}\right) \\
= \quad & \exp\left(w_{0k} + \mathbf{x}^T\boldsymbol{w}_k\right) \exp\left(-\tfrac{1}{2}\mathbf{x}^T\Sigma^{-1}\mathbf{x}\right) \\
\propto \quad & \exp\left(w_{0k} + \mathbf{x}^T\boldsymbol{w}_k\right)
\end{aligned}
$$

by defining $w_{0k} := \log \pi_k - \frac{1}{2}\boldsymbol{\mu}_k^T\Sigma^{-1}\boldsymbol{\mu}_k$ and $\boldsymbol{w}_k := \Sigma^{-1}\boldsymbol{\mu}_k$.

By finally taking the log, we can write our transformed scores as linear:

$$f_k(\mathbf{x}) = w_{0k} + \mathbf{x}^T\boldsymbol{w}_k$$

- The above is a little bit "lax" so lets carefully check
- We left out several (pos) multiplicative constants
- $\exp\left(-\frac{1}{2}\mathbf{x}^T\Sigma^{-1}\mathbf{x}\right)$ contains $\mathbf{x}$ but is the same for all classes
- $\log(at + b)$ is still isotonic for $a > 0$