5 Recall that in GDA we model the joint distribution of (x, y) by the following equations: eqnarray*

p(y) = cases ϕ if $y = 1$

Suppose we have already fit ϕ, μ_0, μ_1, and Σ, and now want to predict y given a new point x. To show that GDA results in a classifier that has a linear decision boundary, show the posterior distribution can be written as equation* $p(y = 1 \mid x; \phi, \mu_0, \mu_1, \Sigma) = 1 1 + \exp(-(\theta^T x + \theta_0))$,