4 1) In general the naive Bayes classifier is not linear, but if the likelihood factor $P(x_i|C)$ are from exponential families, the naive Bayes classifier corresponds to a linear classifier in a particular feature space

we have :

$$P(C=1|X) = \sigma\left(\sum \log \frac{P(x_i|C=1)}{P(x_i|C=0)} + \log \frac{P(C=1)}{P(C=0)}\right)$$

if $P(x_i|C)$ is from an exponential family. We have

$$P(x_i|C) = h_i(x_i) e^{(\mu_{ic}^T \phi_i(x_i) - A_i(\mu_{ic}))}$$

then

$$P(C=1|X) = \sigma(\sum W_i^T \phi_i(x_i) + b)$$

$$w_i = \mu_{i1} - \mu_{i0}$$

$$b = \log \frac{P(C=1)}{P(C=0)} - \sum_i (A_i(\mu_{i1}) - A_i(\mu_{i0}))$$

∵ This is similar to logistic regression in the feature space defined by ϕ.

∴ We get the multinomial logistic regression.

2) $\because$ Logistic regression can be written as:

$$\hat{p} = \frac{1}{1+e^{-\hat{\mu}}} \quad \text{where } \hat{\mu} = \hat{\theta} \cdot x.$$

Since ~~#~~ μ as a ~~predi~~ prediction term is a linear function of x

For logistic regression the decision boundary is linear

$\{x : \hat{p} = 0.5\}$

it's the solution to $\hat{\theta} \cdot x = 0$

$\therefore$ Logistic regression is a linear classifier.