Ex. 6.11

The Gaussian Mixture Model has the form:

$$f(x) = \sum_{m=1}^{M} \alpha_m \phi(x; \mu_m, \Sigma_m)$$

Therefore the likelihood is given by:

$$L(\Theta | X) = \prod_{i=1}^{N} \sum_{m=1}^{M} \alpha_m N(x_i; \mu_m, \Sigma_m)$$

Now consider setting one of the mixture models to be centered on one of the observations e.g. $\mu_K = x_j$ where $K \in 1, \ldots, M$ and $j \in 1, \ldots, N$. Then for that observation mixture combination in the likelihood is given by:

$$\alpha_K (2\pi)^{-K/2} \det(\Sigma_K)^{-1/2} \exp(-\tfrac{1}{2}(x_j - \mu_K) \Sigma_K^{-1} (x_j - \mu_K))$$

$$= \alpha_K (2\pi)^{-K/2} \det(\Sigma_K)^{-1/2}$$

Then as $\Sigma_K \to 0$, this component $\to +\infty$, and therefore the entire likelihood $\to \infty$.

Therefore the likelihood is maximised $\to \infty$ by setting any μ_m to an observation x_i and setting $\Sigma_m \to 0$.