3.b

For each target vector, we have:

$$p(t_i|W,\Sigma) = N(t_i|y(x_i,W),\Sigma) = \frac{1}{\sqrt{(2\pi)^K \det\Sigma}} e^{\frac{-1}{2}(t_i - y(x_i,W))^T \Sigma^{-1}(t_i - y(x_i,W))}$$

So the log-likelihood is given by:

$$\log p(t_i|W,\Sigma) = \log \frac{1}{\sqrt{(2\pi)^K \det\Sigma}} e^{-\frac{1}{2}(t_i - y(x_i,W))^T \Sigma^{-1}(t_i - y(x_i,W))} =$$

$$= (-\frac{1}{2}(t_i - y(x_i,W))^T \Sigma^{-1}(t_i - y(x_i,W))) - \log\sqrt{(2\pi)^K \det\Sigma} =$$

$$= (-\frac{1}{2}(t_i - y(x_i,W))^T \Sigma^{-1}(t_i - y(x_i,W))) - \frac{1}{2}\log((2\pi)^K \det\Sigma) =$$

$$= (-\frac{1}{2}(t_i - y(x_i,W))^T \Sigma^{-1}(t_i - y(x_i,W))) - \frac{K}{2}\log(2\pi) - \frac{1}{2}\log\det\Sigma) =$$

$$= -\frac{1}{2}((t_i - y(x_i,W))^T \Sigma^{-1}(t_i - y(x_i,W)) + K\log 2\pi + \log\det\Sigma)$$

Now that we have the log-likelihood for each target, we can compute the log-likelihood for all N independent observations. It will turn out to be a sum of log terms, since we can multiply the likelihoods for all the independent observations. So we have:

$$p(T|W,\Sigma) = \prod_{i=1}^{N} N(t_i|y(x_i,W),\Sigma)$$

$$\log p(T|W,\Sigma) = \sum_{i=1}^{N} \log N(t_i|y(x_i,W),\Sigma) =$$

And now using the previous results:

$$= \sum_{i=1}^{N} -\frac{1}{2}((t_i - y(x_i,W))^T \Sigma^{-1}(t_i - y(x_i,W)) + K \log 2\pi + \log \det \Sigma) =$$

$$= -\frac{1}{2}(NK \log 2\pi + N \log \det \Sigma + \sum_{i=1}^{N}(t_i - y(x_i,W))^T \Sigma^{-1}(t_i - y(x_i,W)))$$

So the log-likelihood of T is given by $\log p(T|W,\Sigma) = -\frac{1}{2}(NK \log 2\pi + N \log \det \Sigma + \sum_{i=1}^{N}(t_i - y(x_i,W))^T \Sigma^{-1}(t_i - y(x_i,W)))$.