Ex 3.11

We will (conveniently) skip over some of the tricky matrix differenciation here by ~~takin~~ making use of the books step from 3.38 to 3.39.

e.g.

If $RSS(B) = tr[(y - XB)^T (y - XB)]$ $(*)$

then $X^T(y - X\hat{B}) = 0$

and $\hat{B} = (X^T X)^{-1} X^T y$

Now we wish to estimate $\hat{B}$ when

$$RSS(B, \Sigma) = tr[(y - XB)\Sigma^{-1}(y - XB)^T]$$

(using $tr(A^T B) = tr(AB^T)$)

Now, since Σ^{-1} is positive definite and symmetric we can rewrite it as $\Sigma^{-1} = SS^T$ where S is also positive definite and symetric.

$$\begin{aligned}
\Rightarrow RSS &= tr[(y - XB)SS^T(y - XB)^T] \\
&= tr[(yS - XBS)S^T(y^T - B^T X^T)] \\
&= tr[(yS - XBS)(S^T y^T - S^T B^T X^T)] \\
&= tr[(yS - XBS)((yS)^T - (XBS)^T)] \\
&= tr[(yS - XBS)(yS - XBS)^T] \\
&= tr[(yS - XBS)^T(yS - XBS)]
\end{aligned}$$

Now using $(*)$

$$\begin{aligned}
X^T(yS - X\hat{B}S) &= 0 \\
X^T yS &= X^T X\hat{B}S \\
(X^T X)^{-1} X^T yS &= \hat{B}S \\
(X^T X)^{-1} X^T y &= \hat{B}
\end{aligned}$$

as required