2. **Linear Regression With Differing Variances (Adapted from Stanford CS 229)**

Suppose we have a training set $\{\mathbf{x}_n\}$ of N independent examples with corresponding target values $\{t_n\}$, but in which the t_n's were observed with differing variances. Specifically, suppose that

$$p(t_n|\mathbf{x}_n, \mathbf{w}) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{(t_n - \mathbf{w}^\intercal \mathbf{x}_n)^2}{2\sigma_n^2}\right)$$

I.e., t_n has mean $\mathbf{w}^\intercal \mathbf{x}_n$ and variance σ_n^2 (where the σ_n's are fixed, known constants). Show that finding the maximum likelihood estimate of $\mathbf{w}$ reduces to solving a weighted linear regression problem. State clearly what the r_n's (weighting factors) are in terms of the σ_n's.

$$\begin{aligned}
\arg\max_{\mathbf{w}} \prod_{n=1}^{N} p(y_n|\mathbf{x_n}, \mathbf{w}) &= \arg\max_{\mathbf{w}} \sum_{n=1}^{N} \log p(y_n|\mathbf{x_n}, \mathbf{w}) \\
&= \arg\max_{\mathbf{w}} \sum_{n=1}^{N} \left(\log \frac{1}{\sqrt{2\pi}\sigma_n} - \frac{(y_n - \mathbf{w}^\intercal \mathbf{x}_n)^2}{2\sigma_n^2}\right) \\
&= \arg\max_{\mathbf{w}} - \sum_{n=1}^{N} \frac{(y_n - \mathbf{w}^\intercal \mathbf{x}_n)^2}{2\sigma_n^2} \\
&= \arg\min_{\mathbf{w}} \frac{1}{2} \sum_{n=1}^{N} \frac{1}{\sigma_n^2} (y_n - \mathbf{w}^\intercal \mathbf{x}_n)^2 \\
&= \arg\min_{\mathbf{w}} \frac{1}{2} \sum_{n=1}^{N} r_n (y_n - \mathbf{w}^\intercal \mathbf{x}_n)^2
\end{aligned}$$

where in the last step, we substituted $r_n = \dfrac{1}{\sigma_n^2}$ to get the linear regression form.