the variance of the error term ε_i, denoted by σ_i^2, can be expressed as follows:

$$\sigma_i^2 = E\{\varepsilon_i^2\} - (E\{\varepsilon_i\})^2 \tag{11.14}$$

Since $E\{\varepsilon_i\} = 0$ according to the regression model, we obtain:

$$\sigma_i^2 = E\{\varepsilon_i^2\} \tag{11.15}$$

Hence, the squared residual e_i^2 is an estimator of σ_i^2. Furthermore, the absolute residual $|e_i|$ is an estimator of the standard deviation σ_i, since $\sigma_i = |\sqrt{\sigma_i^2}|$.

We can therefore estimate the variance function describing the relation of σ_i^2 to relevant predictor variables by first fitting the regression model using unweighted least squares and then regressing the squared residuals e_i^2 against the appropriate predictor variables. Alternatively, we can estimate the standard deviation function describing the relation of σ_i to relevant predictor variables by regressing the absolute residuals $|e_i|$ obtained from fitting the regression model using unweighted least squares against the appropriate predictor variables. If there are any outliers in the data, it is generally advisable to estimate the standard deviation function rather than the variance function, because regressing absolute residuals is less affected by outliers than regressing squared residuals. Reference 11.1 provides a detailed discussion of the issues encountered in estimating variance and standard deviation functions.

We illustrate the use of some possible variance and standard deviation functions:

1. A residual plot against X_1 exhibits a megaphone shape. Regress the absolute residuals against X_1.
2. A residual plot against $\hat{Y}$ exhibits a megaphone shape. Regress the absolute residuals against $\hat{Y}$.
3. A plot of the squared residuals against X_3 exhibits an upward tendency. Regress the squared residuals against X_3.
4. A plot of the residuals against X_2 suggests that the variance increases rapidly with increases in X_2 up to a point and then increases more slowly. Regress the absolute residuals against X_2 and X_2^2.

After the variance function or the standard deviation function is estimated, the fitted values from this function are used to obtain the estimated weights:

$$w_i = \frac{1}{(\hat{s}_i)^2} \quad \text{where } \hat{s}_i \text{ is fitted value from standard deviation function} \tag{11.16a}$$

$$w_i = \frac{1}{\hat{v}_i} \quad \text{where } \hat{v}_i \text{ is fitted value from variance function} \tag{11.16b}$$

The estimated weights are then placed in the weight matrix $\mathbf{W}$ in (11.7) and the estimated regression coefficients are obtained by (11.9), as follows:

$$\mathbf{b}_w = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{Y} \tag{11.17}$$

The weighted error mean square MSE_w may be viewed here as an estimator of the proportionality constant k in (11.11). If the modeling of the variance or standard deviation function is done well, the proportionality constant will be near 1 and MSE_w should then be near 1.