The Aitken Model

\[ y = X \beta + \xi, \quad \xi \sim (0, \sigma^2 V). \]

Identical to the Gauss-Markov Linear Model except that \( \text{Var}(\xi) = \sigma^2 V \) instead of \( \sigma^2 I \).

V is assumed to be a known nonsingular Variance matrix.
The Normal Theory Aitken Model adds an assumption of normality:

\[ \varepsilon \sim N(0, \sigma^2V). \]

By the Spectral Decomposition Theorem, there exists a nonsingular symmetric matrix \( V^{1/2} \) such that \( V^{1/2}V^{1/2} = V \).

Using \( V^{-1/2} \) to denote \( (V^{1/2})^{-1} \), we have
\[ V^{-\frac{1}{2}} y = V^{-\frac{1}{2}} X \beta + V^{-\frac{1}{2}} \xi. \]

Let \( \xi = V^{-\frac{1}{2}} y \), \( W = V^{-\frac{1}{2}} X \), and \( \delta = V^{\frac{1}{2}} \xi \).

Note that the equation above becomes
\[ \xi = W \beta + \delta, \] where \( \delta \sim (0, \sigma^2 I) \)

because \[ \text{Var}(\xi) = \text{Var}(V^{-\frac{1}{2}} \xi) = V^{-\frac{1}{2}} \sigma^2 V V^{-\frac{1}{2}} = \sigma^2 V^{-\frac{1}{2}} V^{\frac{1}{2}} V^{\frac{1}{2}} V^{-\frac{1}{2}} = \sigma^2 I. \]
Thus, after transformation, we are back to the Gauss-Markov model we are familiar with.

Suppose we want to estimate $E(y)$.

$$E(y) = E(V^{\frac{1}{2}} V^{-\frac{1}{2}} y) = V^{\frac{1}{2}} E(V^{-\frac{1}{2}} y)$$

$$= V^{\frac{1}{2}} E(z).$$
We already know that the best estimate of $E(\bar{z})$ is

\[
\hat{\bar{z}} = P_\omega \bar{z} = w (w'w)^{-1} w' \bar{z} \\
= \sqrt{V} X ( (\sqrt{V}x)' \sqrt{V} X )^{-1} (\sqrt{V}x)' \sqrt{V} X \\
= \sqrt{V} X ( X' \sqrt{V}^{-1} V^{-\frac{1}{2}} X )^{-1} X' \sqrt{V}^{-\frac{1}{2}} V^{-\frac{1}{2}} X \\
= \sqrt{V} X ( X' V^{-1} X )^{-1} X' V^{-1} X .
\]

Thus, to estimate $E (\bar{y}) = \sqrt{V} E (\bar{z})$, we should use $X (X' V^{-1} X )^{-1} X' V^{-1} X$. 
Likewise, if $C\beta$ is estimable, we know the BLUE is the ordinary least squares (OLS) estimator

$$C(w'w)^{-1}w'y = C(x'V^{-\frac{1}{2}}V^{\frac{1}{2}}x)^{-1}x'V^{-\frac{1}{2}}V^{\frac{1}{2}}y = C(x'V^{-1}x)^{-1}x'V^{-1}y.$$ 

$$C(x'V^{-1}x)^{-1}x'V^{-1}y \equiv C_{\hat{\beta}_v}$$ is called a Generalized Least Squares (GLS) estimator.
\( \hat{\beta}_v = (X'V^{-1}X)^{-1}X'V^{-1}y \) is a solution to the Aitken Equations:

\[
X'V^{-1}X\hat{\beta} = X'V^{-1}y
\]

which follow from the Normal Equations

\[
W'W\hat{\beta} = W'y \iff X'V^{-\frac{1}{2}}V^{-\frac{1}{2}}X\hat{\beta} = X'V^{-\frac{1}{2}}V^{-\frac{1}{2}}y
\]

\[
\iff X'V^{-1}X\hat{\beta} = X'V^{-1}y .
\]
Recall that solving the Normal Equations is equivalent to minimizing 
\[(z - Wb)'(z - Wb)\] over \(b \in \mathbb{R}^p\).

Now 
\[(z - Wb)'(z - Wb)\]

\[= (V^{-\frac{1}{2}}y - V^{-\frac{1}{2}}XB)'(V^{-\frac{1}{2}}y - V^{-\frac{1}{2}}XB)\]

\[= (y - XB)'V^{-1}(y - XB).\]
Thus, $\hat{\beta}_v = (x'v^{-1}x)^{-1}x'v^{-1}y$ is a solution to this Generalized Least Squares problem.

When $V$ is diagonal, the term "Weighted Least Squares" (WLS) might be used instead of GLS.
An unbiased estimator of $\sigma^2$ is

$$\frac{\overline{z}'(I-P_W)\overline{z}}{n - \text{rank}(W)} = \frac{|| (I-P_W)\overline{z} ||^2}{n - \text{rank}(W)}$$

$$= \frac{|| (I-W(W'W)^{-1}W)\overline{z} ||^2}{n - \text{rank}(W)}$$

$$= \frac{|| (I - V^{-\frac{1}{2}}X (X'V^{-1}X)^{-1}X'V^{-\frac{1}{2}})V^{-\frac{1}{2}}Y ||^2}{n - \text{rank}(V^{-\frac{1}{2}}X)}$$
$$= \frac{||V^{-\frac{1}{2}} \hat{y} - V^{-\frac{1}{2}} X (X' V^{-1} X)^{-1} X' V^{-1} y||^2}{n - \text{rank}(X)}$$

$$= \frac{||V^{-\frac{1}{2}} (\hat{y} - X \hat{\beta}_v) ||^2}{n - k}$$

$$= \frac{||V^{-\frac{1}{2}} (\hat{y} - X \hat{\beta}_v) ||^2}{(n-k)}$$

$$= (\hat{y} - X \hat{\beta}_v)' V^{-1} (\hat{y} - X \hat{\beta}_v) / (n-k)$$

$$= \frac{\hat{\sigma}_v^2}{\sigma_v^2}.$$
Under the Normal Theory Aitken Model, we can play the same game as above to convert known formulas in terms of $Z$ and $W$ to formulas in terms of $Y$ and $X$ to allow inference about $\beta$ when $Y = X\beta + \varepsilon$, $\varepsilon \sim N(0, \sigma^2 I)$. 