3 Sequential Least Squares

Consider the Generalized Least Square averaging formula applied to the Linear Model $Y = \mathbf{X}\beta + \epsilon$. The covariance matrix for the measurements y is given by $\mathbf{Cov} = E(\epsilon\epsilon^T)$. The Least Squares estimate for β is given by

$$\hat{\beta}_N = \big(\sum_{i=1}^{N} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} \mathbf{X}_i\big)^{-1} \big(\sum_{i=1}^{N} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} Y_i\big)$$

Let $\mathbf{A}_N = \big(\sum_{i=1}^{N} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} \mathbf{X}_i\big)^{-1}$ and we can write

$$\hat{\beta}_N = \mathbf{A}_N \big(\sum_{i=1}^{N} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} Y_i\big)$$

Now consider what happens when we add a new measurement to our average. Using all the data again in a batch estimate we have

$$\hat{\beta}_{N+1} = \mathbf{A}_{N+1} \big(\sum_{i=1}^{N+1} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} Y_i\big),$$

with $\mathbf{A}_{N+1} = \big(\sum_{i=1}^{N+1} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} \mathbf{X}_i\big)^{-1}$ Let us break off the last measurement $i = N+1$ and consider the separate items:

$$\hat{\beta}_{N+1} = \big(\sum_{i=1}^{N} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} \mathbf{X}_i + \mathbf{X}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} \mathbf{X}_{N+1}\big)^{-1} \big(\sum_{i=1}^{N} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} Y_i + \mathbf{X}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} Y_{N+1}\big) \tag{3}$$

Comparing the above expressions we can also express the obvious relationship between $\mathbf{A}_N$ and $\mathbf{A}_{N+1}$ as

$$\mathbf{A}_{N+1}^{-1} = \mathbf{A}_N^{-1} + \mathbf{X}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} \mathbf{X}_{N+1}.$$

In terms of the above matrices we can write the previous estimate with N data points:

$$\sum_{i=1}^{N} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} Y_i = \mathbf{A}_N^{-1} \hat{\beta}_N = \big(\mathbf{A}_{N+1}^{-1} - \mathbf{X}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} \mathbf{X}_{N+1}\big) \hat{\beta}_N$$

The Sherman-Morrison-Woodbury formula for inverses takes the form;

$$(\mathbf{A}_N^{-1} + \mathbf{U}\mathbf{B}^{-1}\mathbf{V}^T)^{-1} = \mathbf{A}_N - \mathbf{A}_N \mathbf{U}(\mathbf{B} + \mathbf{V}^T \mathbf{A}_N \mathbf{U})^{-1} \mathbf{V}^T \mathbf{A}_N$$

Let $\mathbf{A}_N$ be defined as above, $\mathbf{B} = \mathbf{Cov}_{N+1}$, $\mathbf{U} = \mathbf{X}_{N+1}^T$ and $\mathbf{V}^T = \mathbf{X}_{N+1}$ Inserting these expressions into the first term on the RHS of Eq.(3) we obtain the update to the fitted Covariance Matrix:

$$\begin{aligned}
\mathbf{A}_{N+1} &= \big(\sum_{i=1}^{N} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} \mathbf{X}_i + \mathbf{X}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} \mathbf{X}_{N+1}\big)^{-1} \\
&= \mathbf{A}_N - \mathbf{A}_N \mathbf{X}_{N+1}^T (\mathbf{Cov}_{N+1} + \mathbf{X}_{N+1} \mathbf{A}_N \mathbf{X}_{N+1}^T)^{-1} \mathbf{X}_{N+1} \mathbf{A}_N
\end{aligned}$$

We also define a Gain Matrix:

$$\mathbf{K}_{N+1} = \mathbf{A}_{N+1} \mathbf{X}_{N+1}^T \mathbf{Cov}_{N+1}^{-1}$$

We also obtain the update to the fitted parameter state:

$$\begin{aligned}
\hat{\beta}_{N+1} &= \mathbf{A}_{N+1} \big(\sum_{i=1}^{N} \mathbf{X}_i^T \mathbf{Cov}_i^{-1} Y_i + \mathbf{X}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} Y_{N+1}\big) \\
&= \mathbf{A}_{N+1} \big[\big(\mathbf{A}_{N+1}^{-1} - \mathbf{X}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} \mathbf{X}_{N+1}\big) \hat{\beta}_N + \mathbf{X}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} Y_{N+1}\big] \\
&= \hat{\beta}_N + \mathbf{A}_{N+1} \mathbf{X}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} \big(Y_{N+1} - \mathbf{X}_{N+1} \hat{\beta}_N\big) \\
&= \hat{\beta}_N + \mathbf{K}_{N+1} \big(Y_{N+1} - \mathbf{X}_{N+1} \hat{\beta}_N\big)
\end{aligned}$$