

1 Sequential Least Squares

Consider the Generalized Least Square averaging formula applied to the Linear Model $Y = \mathbf{H}\beta + \epsilon$. 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

$$\hat{\beta}_N = \left(\sum_{i=1}^N \mathbf{H}_i^T \mathbf{Cov}_i^{-1} \mathbf{H}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{H}_i^T \mathbf{Cov}_i^{-1} Y_i \right) \quad (1)$$

Let $\mathbf{A}_N = \left(\sum_{i=1}^N \mathbf{H}_i^T \mathbf{Cov}_i^{-1} \mathbf{H}_i \right)^{-1}$ be the coefficient matrix on the RHS of Eq. (1). \mathbf{A}_N is the estimated Covariance of $\hat{\beta}_N$. We can write $\hat{\beta}_N = \mathbf{A}_N \left(\sum_{i=1}^N \mathbf{H}_i^T \mathbf{Cov}_i^{-1} Y_i \right)$. Now consider what happens when we add a new measurement. Using all the data again in a batch estimate we have $\hat{\beta}_{N+1} = \mathbf{A}_{N+1} \left(\sum_{i=1}^{N+1} \mathbf{H}_i^T \mathbf{Cov}_i^{-1} Y_i \right)$, with $\mathbf{A}_{N+1} = \left(\sum_{i=1}^{N+1} \mathbf{H}_i^T \mathbf{Cov}_i^{-1} \mathbf{H}_i \right)^{-1}$. Let us break off the last measurement $i = N+1$ and consider the separate items:

$$\hat{\beta}_{N+1} = \left(\sum_{i=1}^N \mathbf{H}_i^T \mathbf{Cov}_i^{-1} \mathbf{H}_i + \mathbf{H}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} \mathbf{H}_{N+1} \right)^{-1} \left(\sum_{i=1}^N \mathbf{H}_i^T \mathbf{Cov}_i^{-1} Y_i + \mathbf{H}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} Y_{N+1} \right) \quad (2)$$

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

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

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

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

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

$$(\mathbf{A}_N^{-1} + \mathbf{U} \mathbf{B}^{-1} \mathbf{V}^T)^{-1} = \mathbf{A}_N - \mathbf{A}_N \mathbf{U} (\mathbf{B} + \mathbf{V}^T \mathbf{A}_N \mathbf{U})^{-1} \mathbf{V}^T \mathbf{A}_N \quad (3)$$

Let \mathbf{A}_N be defined as above, $\mathbf{B} = \mathbf{Cov}_{N+1}$, $\mathbf{U} = \mathbf{H}_{N+1}^T$ and $\mathbf{V}^T = \mathbf{H}_{N+1}$. Inserting these expressions and using Eq. (3) into the first term on the RHS of Eq.(2) we obtain the update to the fitted Covariance Matrix:

$$\begin{aligned} \mathbf{A}_{N+1} &= \left(\sum_{i=1}^N \mathbf{H}_i^T \mathbf{Cov}_i^{-1} \mathbf{H}_i + \mathbf{H}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} \mathbf{H}_{N+1} \right)^{-1} \\ &= \mathbf{A}_N - \mathbf{A}_N \mathbf{H}_{N+1}^T (\mathbf{Cov}_{N+1} + \mathbf{H}_{N+1} \mathbf{A}_N \mathbf{H}_{N+1}^T)^{-1} \mathbf{H}_{N+1} \mathbf{A}_N \end{aligned} \quad (4)$$

We also obtain the update to the fitted parameter state:

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

Using Eq.(4) we also can express

$$\begin{aligned} \mathbf{A}_{N+1} \mathbf{H}_{N+1}^T &= [\mathbf{A}_N - \mathbf{A}_N \mathbf{H}_{N+1}^T (\mathbf{Cov}_{N+1} + \mathbf{H}_{N+1} \mathbf{A}_N \mathbf{H}_{N+1}^T)^{-1} \mathbf{H}_{N+1} \mathbf{A}_N] \mathbf{H}_{N+1}^T \\ &= \mathbf{A}_N \mathbf{H}_{N+1}^T [\mathbf{I}_N - (\mathbf{Cov}_{N+1} + \mathbf{H}_{N+1} \mathbf{A}_N \mathbf{H}_{N+1}^T)^{-1} \mathbf{H}_{N+1} \mathbf{A}_N \mathbf{H}_{N+1}^T] \\ &= \mathbf{A}_N \mathbf{H}_{N+1}^T (\mathbf{Cov}_{N+1} + \mathbf{H}_{N+1} \mathbf{A}_N \mathbf{H}_{N+1}^T)^{-1} [(\mathbf{Cov}_{N+1} + \mathbf{H}_{N+1} \mathbf{A}_N \mathbf{H}_{N+1}^T) - \mathbf{H}_{N+1} \mathbf{A}_N \mathbf{H}_{N+1}^T] \\ &= \mathbf{A}_N \mathbf{H}_{N+1}^T (\mathbf{Cov}_{N+1} + \mathbf{H}_{N+1} \mathbf{A}_N \mathbf{H}_{N+1}^T)^{-1} \mathbf{Cov}_{N+1} \end{aligned}$$

If we define a Gain Matrix as follows, $\mathbf{K}_{N+1} = \mathbf{A}_{N+1} \mathbf{H}_{N+1}^T \mathbf{Cov}_{N+1}^{-1} = \mathbf{A}_N \mathbf{H}_{N+1}^T (\mathbf{Cov}_{N+1} + \mathbf{H}_{N+1} \mathbf{A}_N \mathbf{H}_{N+1}^T)^{-1}$, we can also write the update formulae as:

$$\mathbf{A}_{N+1} = [\mathbf{I}_N - \mathbf{K}_{N+1} \mathbf{H}_{N+1}] \mathbf{A}_N \quad \text{and} \quad \hat{\beta}_{N+1} = \hat{\beta}_N + \mathbf{K}_{N+1} (Y_{N+1} - \mathbf{H}_{N+1} \hat{\beta}_N).$$

References

- [1] Paul Avery, “Applied Fitting Theory I: Generalized Least Squares Theory, CBX-91-72”
- [2] O.I. Dahl, T.B. Day, F.T. Solmitz, N. L. Gould, “Group A Programming Note No. P-126”
- [3] Stephen P. Smith, “The Cholesky Decomposition and Its Derivatives”
- [4] Escobal, “Methods of Orbit Determination”
- [5] Brouwer and Clemence, “Celestial Mechanics”
- [6] Tapley, Schutz, Born, “Statistical Orbit Determination”
- [7] Griewank, *et al.*, “Computational Differentiation”
- [8] John Georganas, William Kovalcik, “Angle Sensor Tracking Concepts”
- [9] John Georganas, “Converted Specific Angular Momentum Measurements Estimator”
- [10] Elizabeth M. Keil, “Kalman Filter Implementation to Determining Orbit and Attitude of a Satellite in a Molniya Orbit”
- [11] David R. Bizup, Donald E. Brown, “The Over-Extended Kalman Filter – Don’t Use It!”
- [12] Roger R. Bate, Donald D. Mueller, Jerry E. White, “Fundamentals of Astrodynamics”, Dover 1971, pp. 117-122
- [13] P.O. Bell, “Orbit Determination by Angular Measurements”
- [14] A.J.Claus, R.B. Blackman, E.G. Halline, W.C. Riggway, III, “Orbit Determination and Prediction and Computer Programs”
- [15] Alvin C. Rencher and G. Bruce Schaalje, “Linear Models in Statistics”