The Agee-Turner Factorization Update Algorithm

Eun Hwan Shin

October 18, 2022

1 Scope

This document fills in between equations in Section III.4 of the book "Gerald J. Bierman. (1977). Factorization Methods for Discrete Sequential Estimation. Academic Press, Inc., New York."

2 Description

Let $\mathbf{P}$ be an $n \times n$ symmetric positive definite matrix factorized as $\mathbf{P} = \mathbf{U}\mathbf{D}\mathbf{U}^T$, where $\mathbf{U}$ and $\mathbf{D}$ are a unit upper-triangular matrix and a diagonal matrix, respectively. The Agee-Turner update algorithm seeks to find the updated $\bar{\mathbf{U}}$ and $\bar{\mathbf{D}}$ such that

$$\bar{\mathbf{U}}\bar{\mathbf{D}}\bar{\mathbf{U}}^T = \mathbf{U}\mathbf{D}\mathbf{U}^T + c\mathbf{a}\mathbf{a}^T$$

where $\mathbf{a}$ is a vector and c is a scalar.

Let $\mathbf{x}$ be an arbitrary vector of size n and $\mathbf{v} = \mathbf{U}^T\mathbf{x}$. The last row of $\mathbf{U}^T$ is

$$\begin{bmatrix} U(1,n) & U(2,n) & \cdots & U(n-1,n) & 1 \end{bmatrix}$$

Let d_i's denote the diagonal elements of $\mathbf{D}$. Then,

$$v_n = x_n + \sum_{j=1}^{n-1} U(j,n)x_j \tag{1}$$

$$d_n v_n^2 = d_n x_n^2 + 2d_n x_n \sum_{j=1}^{n-1} U(j,n)x_j + d_n \left(\sum_{j=1}^{n-1} U(j,n)x_j \right)^2 \tag{2}$$

$$\begin{aligned} c(\mathbf{a}^T\mathbf{x})^2 &= c\left(a_n x_n + \sum_{j=1}^{n-1} a_j x_j \right)^2 \\ &= c a_n^2 x_n^2 + 2c a_n x_n \sum_{j=1}^{n-1} a_j x_j + c\left(\sum_{j=1}^{n-1} a_j x_j \right)^2 \end{aligned} \tag{3}$$

Thus, we can obtain Eqs. (4.6) and (4.7) of the book as follows:

$$\begin{aligned}
\mathbf{x}^T\bar{\mathbf{P}}\mathbf{x} &= \sum_{j=1}^{n} d_j v_j^2 + c(\mathbf{a}^T\mathbf{x})^2 && (4)\\
&= \sum_{j=1}^{n-1} d_j v_j^2 + d_n v_n^2 + c(\mathbf{a}^T\mathbf{x})^2\\
&= \sum_{j=1}^{n-1} d_j v_j^2 + d_n x_n^2 + 2d_n x_n \sum_{j=1}^{n-1} U(j,n)x_j + d_n \left(\sum_{j=1}^{n-1} U(j,n)x_j\right)^2\\
&\quad + c a_n^2 x_n^2 + 2c a_n x_n \sum_{j=1}^{n-1} a_j x_j + c\left(\sum_{j=1}^{n-1} a_j x_j\right)^2\\
&= \sum_{j=1}^{n-1} d_j v_j^2 + (d_n + c a_n^2)x_n^2 + 2x_n \sum_{j=1}^{n-1} \left(d_n U(j,n) + c a_n a_j\right) x_j\\
&\quad + d_n \left(\sum_{j=1}^{n-1} U(j,n)x_j\right)^2 + c\left(\sum_{j=1}^{n-1} a_j x_j\right)^2 && (5)
\end{aligned}$$

Let $\bar{d}_n = d_n + c a_n^2$ and

$$\begin{aligned}
\bar{\mathbf{w}}^T &= \begin{bmatrix} U(1,n) & U(2,n) & \cdots & U(n-1,n) \end{bmatrix}\\
\bar{\mathbf{a}}^T &= \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} \end{bmatrix}\\
\bar{\mathbf{x}}^T &= \begin{bmatrix} x_1 & x_2 & \cdots & x_{n-1} \end{bmatrix}
\end{aligned}$$

Then,

$$\bar{\mathbf{w}}^T\bar{\mathbf{x}} = \sum_{j=1}^{n-1} U(j,n)x_j, \quad \bar{\mathbf{a}}^T\bar{\mathbf{x}} = \sum_{j=1}^{n-1} a_j x_j$$

and

$$(d_n\bar{\mathbf{w}} + c a_n \bar{\mathbf{a}})^T\bar{\mathbf{x}} = \sum_{j=1}^{n-1} \left(d_n U(j,n) + c a_n a_j\right) x_j$$

Eq. (5) can be rewritten as

$$\mathbf{x}^T\bar{\mathbf{P}}\mathbf{x} = \sum_{j=1}^{n-1} d_j v_j^2 + \bar{d}_n x_n^2 + 2x_n(d_n\bar{\mathbf{w}} + c a_n\bar{\mathbf{a}})^T\bar{\mathbf{x}} + d_n(\bar{\mathbf{w}}^T\bar{\mathbf{x}})^2 + c(\bar{\mathbf{a}}^T\bar{\mathbf{x}})^2 \qquad (6)$$

Using

$$\begin{aligned}
&\bar{d}_n x_n^2 + 2x_n(d_n\bar{\mathbf{w}} + c a_n\bar{\mathbf{a}})^T\bar{\mathbf{x}} = \bar{d}_n\left[x_n^2 + 2x_n(d_n\bar{\mathbf{w}} + c a_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}/\bar{d}_n\right]\\
&\quad = \bar{d}_n\left[x_n^2 + 2x_n(d_n\bar{\mathbf{w}} + c a_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}/\bar{d}_n + \left\{(d_n\bar{\mathbf{w}} + c a_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}/\bar{d}_n\right\}^2\right]\\
&\qquad -\bar{d}_n\left\{(d_n\bar{\mathbf{w}} + c a_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}/\bar{d}_n\right\}^2\\
&\quad = \bar{d}_n\left[x_n + (d_n\bar{\mathbf{w}} + c a_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}/\bar{d}_n\right]^2 - \left[(d_n\bar{\mathbf{w}} + c a_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}\right]^2/\bar{d}_n
\end{aligned}$$

we can obtain Eq. (4.8) of the book:

$$\begin{aligned}
\mathbf{x}^T\bar{\mathbf{P}}\mathbf{x} &= \sum_{j=1}^{n-1} d_j v_j^2 + \bar{d}_n \left[x_n + (d_n\bar{\mathbf{w}} + ca_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}/\bar{d}_n\right]^2 \\
&\quad + d_n(\bar{\mathbf{w}}^T\bar{\mathbf{x}})^2 + c(\bar{\mathbf{a}}^T\bar{\mathbf{x}})^2 - \left[(d_n\bar{\mathbf{w}} + ca_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}\right]^2/\bar{d}_n \qquad (7)
\end{aligned}$$

Since

$$\begin{aligned}
d_n(\bar{\mathbf{w}}^T\bar{\mathbf{x}})^2 &= d_n(\bar{\mathbf{x}}^T\bar{\mathbf{w}})(\bar{\mathbf{w}}^T\bar{\mathbf{x}}) = d_n\bar{\mathbf{x}}^T(\bar{\mathbf{w}}\bar{\mathbf{w}}^T)\bar{\mathbf{x}} \\
c(\bar{\mathbf{a}}^T\bar{\mathbf{x}})^2 &= c(\bar{\mathbf{x}}^T\bar{\mathbf{a}})(\bar{\mathbf{a}}^T\bar{\mathbf{x}}) = c\bar{\mathbf{x}}^T(\bar{\mathbf{a}}\bar{\mathbf{a}}^T)\bar{\mathbf{x}} \\
\left[(d_n\bar{\mathbf{w}} + ca_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}\right]^2/\bar{d}_n &= \bar{\mathbf{x}}^T(d_n\bar{\mathbf{w}} + ca_n\bar{\mathbf{a}})(d_n\bar{\mathbf{w}} + ca_n\bar{\mathbf{a}})^T\mathbf{x}/\bar{d}_n \\
&= \bar{\mathbf{x}}^T\left[(d_n^2/\bar{d}_n)\bar{\mathbf{w}}\bar{\mathbf{w}}^T + (2cd_na_n/\bar{d}_n)\bar{\mathbf{w}}\bar{\mathbf{a}}^T + (c^2a_n^2/\bar{d}_n)\bar{\mathbf{a}}\bar{\mathbf{a}}^T\right]\bar{\mathbf{x}}
\end{aligned}$$

we can write

$$\begin{aligned}
& d_n(\bar{\mathbf{w}}^T\bar{\mathbf{x}})^2 + c(\bar{\mathbf{a}}^T\bar{\mathbf{x}})^2 - \left[(d_n\bar{\mathbf{w}} + ca_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}\right]^2/\bar{d}_n \\
&= \bar{\mathbf{x}}^T\left[d_n(1 - d_n/\bar{d}_n)\bar{\mathbf{w}}\bar{\mathbf{w}}^T - (2cd_na_n/\bar{d}_n)\bar{\mathbf{w}}\bar{\mathbf{a}}^T + c(1 - ca_n^2/\bar{d}_n)\bar{\mathbf{a}}\bar{\mathbf{a}}^T\right]\bar{\mathbf{x}} \\
&= \bar{\mathbf{x}}^T\left[(d_nca_n^2/\bar{d}_n)\bar{\mathbf{w}}\bar{\mathbf{w}}^T - (2cd_na_n/\bar{d}_n)\bar{\mathbf{w}}\bar{\mathbf{a}}^T + (cd_n/\bar{d}_n)\bar{\mathbf{a}}\bar{\mathbf{a}}^T\right]\bar{\mathbf{x}} \\
&= (cd_n/\bar{d}_n)\bar{\mathbf{x}}^T\left[a_n^2\bar{\mathbf{w}}\bar{\mathbf{w}}^T - 2a_n\bar{\mathbf{w}}\bar{\mathbf{a}}^T + \bar{\mathbf{a}}\bar{\mathbf{a}}^T\right]\bar{\mathbf{x}} \\
&= (cd_n/\bar{d}_n)\bar{\mathbf{x}}^T(\bar{\mathbf{a}} - a_n\bar{\mathbf{w}})(\bar{\mathbf{a}} - a_n\bar{\mathbf{w}})^T\bar{\mathbf{x}} \\
&= (cd_n/\bar{d}_n)\left[(\bar{\mathbf{a}} - a_n\bar{\mathbf{w}})^T\bar{\mathbf{x}}\right]^2
\end{aligned}$$

Then, with

$$y_n = x_n + (d_n\bar{\mathbf{w}} + ca_n\bar{\mathbf{a}})^T\bar{\mathbf{x}}/\bar{d}_n \qquad (8)$$

we can obtain Eq. (4.9) of the book:

$$\mathbf{x}^T\bar{\mathbf{P}}\mathbf{x} = \bar{d}_ny_n^2 + \left[\sum_{j=1}^{n-1} d_jv_j^2 + (cd_n/\bar{d}_n)\left\{(\bar{\mathbf{a}} - a_n\bar{\mathbf{w}})^T\bar{\mathbf{x}}\right\}^2\right] \qquad (9)$$

The terms inside brackets in Eq. (9) are in the same form as in Eq. (4). Thus, in the next stage, it is requited to have

$$\begin{aligned}
\bar{\mathbf{a}} &\leftarrow \bar{\mathbf{a}} - a_n\bar{\mathbf{w}} && (10) \\
\bar{\mathbf{w}} &\leftarrow (d_n\bar{\mathbf{w}} + ca_n\bar{\mathbf{a}})/\bar{d}_n \\
&= (1 - ca_n^2/\bar{d}_n)\bar{\mathbf{w}} + ca_n\bar{\mathbf{a}}/\bar{d}_n \\
&= \bar{\mathbf{w}} + ca_n\left[\bar{\mathbf{a}} - a_n\bar{\mathbf{w}}\right]/\bar{d}_n && (11) \\
c &\leftarrow cd_n/\bar{d}_n && (12)
\end{aligned}$$

where $[\bar{\mathbf{a}} - a_n\bar{\mathbf{w}}]$ in Eq. (11) is an updated $\bar{\mathbf{a}}$ in Eq. (10).

If we apply the same procedure repeatedly while reducing the dimension successively, we can write

$$\mathbf{x}^T\bar{\mathbf{P}}\mathbf{x} = \sum_{j=2}^{n} \bar{d}_jy_j^2 + \bar{d}_1x_1^2$$

Hence, we can identify $\mathbf{y} = \bar{\mathbf{U}}^T \mathbf{x}$ and $x_1 = y_1$. Therefore, the Agee-Turner upadate algorithm can be summarized as Eqs. (4.2)–(4.5) in the book, which is repeated below. For $j = n, n-1, \cdots, 2$,

$$\begin{aligned} \bar{d}_j &= d_j + c_j a_j^2 && (13) \\ a_k &:= a_k - a_j U(k,j), \quad k = 1, \cdots, j-1 && (14) \\ \bar{U}(k,j) &= U(k,j) + c a_j a_k / \bar{d}_j, \quad k = 1, \cdots, j-1 && (15) \\ c &:= c d_j / \bar{d}_j && (16) \end{aligned}$$

Finally, $\bar{d}_1 = d_1 + c_1 a_1^2$.