Computing the Lower Bound Efficiently

We have

$$\|y - Rx\|_2^2 = \|y_{k:m} - R_{k:m,k:m}x_{k:m}\|_2^2 + \|y_{1:k-1} - R_{1:k-1,1:k-1}x_{1:k-1} - R_{1:k-1,k:m}x_{k:m}\|_2^2.$$

Suppose that $U = R^{-1}$ has been computed, then the lower bound on the second term of the above equation is:

$$\sigma_{min}(R_{1:k-1,1:k-1}) \min_{x_{1:k-1}} \|x_{1:k-1} - U_{1:k-1,1:k-1}(y_{1:k-1} - R_{1:k-1,k:m}x_{k:m})\|_2^2.$$

We need to efficiently compute:

$$f^{k-1} \equiv U_{1:k-1,1:k-1}(y_{1:k-1} - R_{1:k-1,k:m}x_{k:m}). \tag{1}$$

For $k = m$ we compute f^{k-1} directly by matrix-vector multiplication. For $k < m$, a recursion is derived to relate f^{k-2} to the already known f^{k-1}.

First split f^{k-1} into 2 parts, $f^{k-1}_{1:k-2}$ and $f^{k-1}_{k-1:m}$ as follows:

$$\begin{aligned}
f^{k-1} &= U_{1:k-1,1:k-1}(y_{1:k-1} - R_{1:k-1,k:m}x_{k:m}) \\
&= \begin{bmatrix} U_{1:k-2,1:k-2} & U_{1:k-2,k-1} \\ 0 & u_{k-1,k-1} \end{bmatrix} \begin{bmatrix} y_{1:k-2} \\ y_{k-1} \end{bmatrix} - \begin{bmatrix} U_{1:k-2,1:k-2} & U_{1:k-2,k-1} \\ 0 & u_{k-1,k-1} \end{bmatrix} \begin{bmatrix} R_{1:k-2,k:m} \\ R_{k-1,k:m} \end{bmatrix} x_{k:m}. \\
&= \begin{bmatrix} U_{1:k-2,1:k-2}y_{1:k-2} + U_{1:k-2,k-1}y_{k-1} \\ u_{k-1,1:k-1}y_{k-1} \end{bmatrix} - \begin{bmatrix} U_{1:k-2,1:k-2}R_{1:k-2,k:m}x_{k:m} + U_{1:k-2,k-1}R_{k-1,k:m}x_{k:m} \\ u_{k-1,k-1}R_{k-1,k:m}x_{k:m} \end{bmatrix}.
\end{aligned}$$

So,

$$f^{k-1}_{1:k-2} = U_{1:k-2,1:k-2}y_{1:k-2} + U_{1:k-2,k-1}y_{k-1} - U_{1:k-2,1:k-2}R_{1:k-2,k:m}x_{k:m} - U_{1:k-2,k-1}R_{k-1,k:m}x_{k:m}. \tag{2}$$

We also know from (1) that

$$\begin{aligned}
f^{k-2} &= U_{1:k-2,1:k-2}y_{1:k-2} - U_{1:k-2,1:k-2}R_{1:k-2,k-1:m}x_{k-1:m} \\
&= U_{1:k-2,1:k-2}y_{1:k-2} - U_{1:k-2,1:k-2}\begin{bmatrix} R_{1:k-2,k-1} & R_{1:k-2,k:m} \end{bmatrix} \begin{bmatrix} x_{k-1} \\ x_{k:m} \end{bmatrix} \\
&= U_{1:k-2,1:k-2}y_{1:k-2} - U_{1:k-2,1:k-2}R_{1:k-2,k:m}x_{k:m} - U_{1:k-2,1:k-2}R_{1:k-2,k-1}x_{k-1}.
\end{aligned} \tag{3}$$

Comparing (2) and (3), we notice that the first 2 terms in (3) also appear in (2) as the first and third term. This gives us the following equation to relate f^{k-1} and f^{k-2}:

$$f^{k-2} = f^{k-1}_{1:k-2} - U_{1:k-2,k-1}y_{k-1} + U_{1:k-2,k-1}R_{k-1,k:m}x_{k:m} - U_{1:k-2,1:k-2}R_{1:k-2,k-1}x_{k-1}. \tag{4}$$

The last term in the above equality involves the matrix-vector multiplication $U_{1:k-2,1:k-2}R_{1:k-2,k-1}$, which can be simplified. In fact, since $UR = I$, it is easy to verify that

$$\begin{bmatrix} U_{1:k-2,1:k-2} & U_{1:k-2,k-1} \end{bmatrix} \begin{bmatrix} R_{1:k-2,k-1} \\ r_{k-1,k-1} \end{bmatrix} = 0.$$

Thus $U_{1:k-2,1:k-2}R_{1:k-2,k-1} = -U_{1:k-2,k-1}r_{k-1,k-1}$. Then from (4) we have

$$\begin{aligned}
f^{k-2} &= f^{k-1}_{1:k-2} - U_{1:k-2,k-1}y_{k-1} + U_{1:k-2,k-1}R_{k-1,k:m}x_{k:m} + U_{1:k-2,k-1}r_{k-1,k-1}x_{k-1} \\
&= f^{k-1}_{1:k-2} - U_{1:k-2,k-1}(y_{k-1} - R_{k-1,k-1:m}x_{k-1:m}).
\end{aligned}$$

Notice that $y_{k-1} - R_{k-1,k-1:m}x_{k-1:m}$ is computed in the search process. So if f^{k-1} is known, the cost for computing f^{k-2} is about $2k$ flops. Note that when $x_{k:m}$ is determined, the lower bound is determined. During the search process, when move from a higher lever to a lower level, we need to compute the lower bound and store the value; when move a lower level to a higher level, we just use the stored values of the lower bounds.