Derivation of L-BFGS

$$
\begin{aligned}
B_1 &= B_0 - \frac{B_0 s_0 s_0^T B_0}{s_0^T B_0 s_0} + \frac{y_0 y_0^T}{y_0^T s_0} \\
&= B_0 - \begin{pmatrix} B_0 s_0 & y_0 \end{pmatrix} \begin{pmatrix} 1/s_0^T B_0 s_0 & 0 \\ 0 & -1/y_0^T s_0 \end{pmatrix} \begin{pmatrix} s_0^T B_0 \\ y_0^T \end{pmatrix} \\
&= B_0 - \begin{pmatrix} B_0 s_0 & y_0 \end{pmatrix} \begin{pmatrix} m_{11} & 0 \\ 0 & m_{33} \end{pmatrix} \begin{pmatrix} s_0^T B_0 \\ y_0^T \end{pmatrix}
\end{aligned}
$$

$$B_2 = B_1 - \frac{B_1 s_1 s_1^T B_1}{s_1^T B_1 s_1} + \frac{y_1 y_1^T}{y_1^T s_1}$$

Can we express B_2 in that form?

$$\begin{pmatrix} B_0 s_0 & B_0 s_1 & y_0 & y_1 \end{pmatrix} M_2 \begin{pmatrix} s_0^T B_0 \\ s_1^T B_0 \\ y_0^T \\ y_1^T \end{pmatrix}$$

M_2 is a 4*4 matrix

First, rewrite B_1 in that form

$$
\begin{aligned}
B_1 &= B_0 - \begin{pmatrix} B_0 s_0 & y_0 \end{pmatrix} \begin{pmatrix} m_{11} & 0 \\ 0 & m_{33} \end{pmatrix} \begin{pmatrix} s_0^T B_0 \\ y_0^T \end{pmatrix} \\
&= \begin{pmatrix} B_0 s_0 & B_0 s_1 & y_0 & y_1 \end{pmatrix} \begin{pmatrix} m_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & m_{33} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} s_0^T B_0 \\ s_1^T B_0 \\ y_0^T \\ y_1^T \end{pmatrix} \\
&= \begin{pmatrix} B_0 s_0 & B_0 s_1 & y_0 & y_1 \end{pmatrix} M_1 \begin{pmatrix} s_0^T B_0 \\ s_1^T B_0 \\ y_0^T \\ y_1^T \end{pmatrix}
\end{aligned}
$$

Second, make that for $B_1 s_1 s_1^T B_1$

$$
\begin{aligned}
B_2 &= B_1 - \frac{B_1 s_1 s_1^T B_1}{s_1^T B_1 s_1} + \frac{y_1 y_1^T}{y_1^T s_1} \\
B_1 s_1 &= B_0 s_1 - m_{11}(u_0^T s_1) u_0 + m_{33}(y_0^T s_1) y_0 \\
&= B_0 s_1 - \hat{m}_{11} B_0 s_0 + \hat{m}_{33} y_0 \\
&= (B_0 s_0 \; B_0 s_1 \; y_0 \; y_1) \begin{pmatrix} \hat{m}_{11} \\ 1 \\ \hat{m}_{33} \\ 0 \end{pmatrix} \\
&= (B_0 s_0 \; B_0 s_1 \; y_0 \; y_1) \, u_1 \\
B_1 s_1 s_1^T B_1 &= (B_0 s_0 \; B_0 s_1 \; y_0 \; y_1) \, u_1 u_1^T \begin{pmatrix} s_0^T B_0 \\ s_1^T B_0 \\ y_0 \\ y_1 \end{pmatrix}
\end{aligned}
$$

Last, put $y_1 y_1^T$ into the form

$$
\begin{aligned}
y_1 y_1^T &= \begin{pmatrix} B_0 s_0 & B_0 s_1 & y_0 & y_1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} s_0^T B_0 \\ s_1^T B_0 \\ y_0^T \\ y_1^T \end{pmatrix} \\
&= \begin{pmatrix} B_0 s_0 & B_0 s_1 & y_0 & y_1 \end{pmatrix} e_4 e_4^T \begin{pmatrix} s_0^T B_0 \\ s_1^T B_0 \\ y_0^T \\ y_1^T \end{pmatrix} \\
B_2 &= B_1 - \frac{B_1 s_1 s_1^T B_1}{s_1^T B_1 s_1} + \frac{y_1 y_1^T}{y_1^T s_1} \\
&= B_0 - \begin{pmatrix} B_0 s_0 & B_0 s_1 & y_0 & y_1 \end{pmatrix} (M_1 - \frac{u_1 u_1^T}{s_1^T B_1 s_1} + \frac{e_4 e_4^T}{y_1^T s_1}) \begin{pmatrix} s_0^T B_0 \\ s_1^T B_0 \\ y_0^T \\ y_1^T \end{pmatrix}
\end{aligned}
$$

More advanced reduction

- Show that

$$\left(M_1 - \frac{u_1 u_1^T}{s_1^T B_1 s_1} + \frac{T_1}{y_1^T y_1}\right) = \begin{pmatrix} S_1^T B_0 S_1 & L_1 \\ L_1^T & -D_1 \end{pmatrix}^{-1}$$

$$S_1 = (s_0 \ s_1), L_1 = \begin{pmatrix} 0 & s_0^T y_1 \\ 0 & 0 \end{pmatrix}, D_1 = \begin{pmatrix} s_0^T y_0 & 0 \\ 0 & s_1^T y_1 \end{pmatrix}$$

- Some linear algebra (LU decomposition + Sherman-Morrison-Woodbury formula)

$$K = C - B^T A^{-1} B$$

$$\begin{pmatrix} A & B \\ B^T & C \end{pmatrix}^{-1} = \begin{pmatrix} (A + BC^{-1}B^T)^{-1} & -A^{-1}BK^{-1} \\ -K^{-1}B^T A^{-1} & K^{-1} \end{pmatrix}$$

Example: partially separable fn

$$f(x) = (x_1 - x_3^2)^2 + (x_2 - x_4^2)^2 + (x_3 - x_2^2)^2 + (x_4 - x_1^2)^2$$

$$\nabla^2 f(x) = H_1 + H_2 + H_3 + H_4$$

$$\mathrm{H}_1 = \begin{pmatrix} 2 & 0 & -4x_3 & 0 \\ 0 & 0 & 0 & 0 \\ -4x_3 & 0 & 12x_3^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \mathrm{H}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -4x_4 \\ 0 & 0 & 0 & 0 \\ 0 & -4x_4 & 0 & 12x_4^2 \end{pmatrix}$$

$$\mathrm{H}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 12x_2^2 & -4x_2 & 0 \\ 0 & -4x_2 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \mathrm{H}_4 = \begin{pmatrix} 12x_1^2 & 0 & 0 & -4x_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -4x_1 & 0 & 0 & 2 \end{pmatrix}$$