Piecewise linear regression for yield curve fitting

Given a set of bond data currently tradable in the market :

$\{T_n, b_n, a, c_n, d_n\}$ for all n belongs to *{1,N}*, where

T_n	is bond maturity
b_n	is current best bid of bond n
a_n	is current best ask of bond n
c_n	is previous close price of bond n
d_n	is dv01 of bond n

As the objective is to fit the yield curve $y(\tau|y_{\tau 0}, y_{\tau 1}, y_{\tau 2,\dots,} y_{\tau M})$, with τ being any tenor point on the yield curve. The curve is parameterized by the yield at a set of predefined time points, that is $\{y_{\tau m}\}$ where m belongs to *{0,M}*. This is a piecewise linear model with M control points, common control points are :

$$\tau_0 = 0$$
$$\tau_1 = 1d$$
$$\tau_2 = 1M$$
$$\tau_3 = 3M$$
$$\tau_4 = 6M$$
$$\tau_5 = 1Y \dots \text{etc}$$

The first step involves converting observed bond prices to bond yields using method like bisection or Newton Raphson (by solving Internal-Rate-Return equation). Alternatively, we can approximate price and yield relation using linear term of Taylor series, as long as the pivot is close enough. Here we take the previous day closing price as the pivot, for bond n :

$$m_n = \tfrac{1}{2}(b_n + a_n) = c_n + d_n y_{Tn}$$

Please note the following different notations :

$\{y_{Tn}\}$	$n = [1,N]$	are observations / data points
$\{y_{\tau m}\}$	$m = [0,M]$	are parameters / control points
$y(\tau\| \dots)$		is the regression of yield curve at τ

The primary objective is to minimize weighted sum of price-error square wr.t. control points :

$$min_{\{y_{\tau m}\}} \textstyle\sum_{n=1}^{N} w_n \left(c_n + d_n y_{Tn} - \frac{1}{2}(b_n + a_n)\right)^2$$

$$\rightarrow \quad min_{\{y_{\tau m}\}} \textstyle\sum_{n=1}^{N} w_n \left(c_n + d_n\left[(1-\mu_{Tn})y_{\tau m} + \mu_{Tn} y_{\tau(m+1)}\right] - \frac{1}{2}(b_n + a_n)\right)^2 \qquad (1)$$

where maturity of bond n lies between 2 tenor points in yield curve $y_{\tau m} \leq y_{Tn} \leq y_{\tau(m+1)}$, weight w_n of a bond price data is inversely proportional to its spread, while μ_{Tn} is for piecewise linear interpolation.

$$w_n = \left(\frac{1}{a_n - b_n}\right)^2 \qquad \mu_{Tn} = \frac{T_n - \tau_m}{\tau_{m+1} - \tau_m}$$

Rewriting *(1)* as matrix form :

$$min_Y (AY - B)^T W (AY - B) \;\; \textit{where W is diagonal matrix of weight} \qquad (2)$$

where A is $N\times(M+1)$ design matrix, Y is $M+1$ parameter vector, B is N output vector.

$$A = \begin{bmatrix} 0 & d_1\frac{\tau_2 - T_1}{\tau_2 - \tau_1} & d_1\frac{T_1 - \tau_1}{\tau_2 - \tau_1} & 0 & \dots & 0 & 0 \\ 0 & 0 & d_2\frac{\tau_3 - T_2}{\tau_3 - \tau_2} & d_2\frac{T_2 - \tau_2}{\tau_3 - \tau_2} & \dots & 0 & 0 \\ d_3\frac{\tau_1 - T_3}{\tau_1 - \tau_0} & d_3\frac{T_3 - \tau_0}{\tau_1 - \tau_0} & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & d_N\frac{\tau_{M+1} - T_N}{\tau_{M+1} - \tau_M} & d_N\frac{T_N - \tau_M}{\tau_{M+1} - \tau_M} \end{bmatrix}$$

$$Y = \begin{bmatrix} y_{\tau 0} \\ y_{\tau 1} \\ y_{\tau 2} \\ \dots \\ y_{\tau M} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{2}(b_1 + a_1) - c_1 \\ \frac{1}{2}(b_2 + a_2) - c_2 \\ \frac{1}{2}(b_3 + a_3) - c_3 \\ \dots \\ \frac{1}{2}(b_N + a_N) - c_N \end{bmatrix}$$

For some unknown reasons, we would like to apply a linear transformation on data space, and hence a corresponding transformation in parameter space (to handle case when all bonds have similar tenor?). The parameter space is transformed from yield to delta yield between 2 tenor points, as a result, a long term bond price observation can contribute to calculate short term tenor point in yield curve.

$$\begin{aligned} & min_Y (AY - B)^T W (AY - B) \\ = \quad & min_Y (AT^{-1}TY - B)^T W (AT^{-1}T - B) \\ = \quad & min_{Y'} (A'Y' - B)^T W (A'Y' - B) \end{aligned} \qquad (3)$$

where $$T = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix} \textit{ is transformation of parameter space}$$

$$A' = AT^{-1} = A\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \qquad Y' = TY = \begin{bmatrix} y_{\tau 0} \\ y_{\tau 1} - y_{\tau 0} \\ y_{\tau 2} - y_{\tau 1} \\ \dots \\ y_{\tau M} - y_{\tau (M-1)} \end{bmatrix}$$

Finally we introduce regularization term. It becomes minimization of *J(Y')* :

$$J(Y') = (A'Y' - B)^T W (A'Y' - B) + \lambda {Y'}^T Y' \qquad (4)$$

$$\frac{dJ}{dY'} = 2{A'}^T W\left(A'{Y'}_{opt} - B\right) + 2\lambda {Y'}_{opt} = 0$$

$${A'}^T W A' {Y'}_{opt} + \lambda {Y'}_{opt} - {A'}^T W B = 0$$

$${Y'}_{opt} = \left({A'}^T W A' + \lambda I\right)^{-1} \left({A'}^T W B\right)$$

The yield curve adjusted mid price $m_{n,adj}$ of bond n can be found from the regression result :

$${A'}_{n,adj} {Y'}_{opt} = B_{n,adj} = m_{n,adj} - c_n$$

$$m_{n,adj} = {A'}_{n,adj} {Y'}_{opt} + c_n \qquad \textit{where } {A'}_{n,adj} = \begin{bmatrix} 0 & d_n\frac{\tau_2 - T_n}{\tau_2 - \tau_1} & d_n\frac{T_n - \tau_1}{\tau_2 - \tau_1} & 0 & \dots & 0 \end{bmatrix}$$

assuming that $\tau_1 \le T_n \le \tau_2$