<u>Steps</u> you have an overdetermined system

1. your basic equation is

$$\text{error}_i + A X_i^4 + B X_i^7 + C x_i + D y_i = R_i \qquad i = 1 \text{ to } m$$

2. This can be written in matrix form as

$$[R] = [X][\beta] + [\varepsilon] \leftarrow \text{error}$$

where
$$[R] = \begin{bmatrix} F \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad [\beta] = \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} \qquad [X] = \begin{bmatrix} X_1^4 & X_1^7 & X_1 & Y_1 \\ X_2^4 & X_2^7 & X_2 & Y_2 \\ \vdots & & & \\ X_m^4 & X_m^7 & X_m & Y_m \end{bmatrix} \qquad [\varepsilon] = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix}$$

errors

3. Solve for errors

$$\Longrightarrow [\varepsilon] = [R] - [X][\beta]$$

4. Square the errors

$$\Longrightarrow [\varepsilon]^T[\varepsilon] = \left([R] - [X][\beta]\right)^T \left([R] - [X][\beta]\right)$$
$$= \left([R]^T - [\beta]^T[X]^T\right)\left([R] - [X][\beta]\right)$$
$$= [R]^T[R] - [R]^T[X][\beta] - [\beta]^T[X]^T[R] + [\beta]^T[X]^T[X][\beta]$$
$$= [R]^T[R] - 2[\beta]^T[X]^T[R] + [\beta]^T[X]^T[X][\beta]$$

4. Minimize the squared Errors by taking

$$\frac{\partial\left([\varepsilon]^T[\varepsilon]\right)}{\partial[\beta]} = [0]$$

and replacing $[\beta]$ with $[b]$ because now we call $[b]$ the best estimates for $[\beta]$ in a least square sense.

$\Longrightarrow$ 4. Continued

$[b]$ is basically the best estimate for $[\beta] = \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix}$

$$\frac{\partial\left([\varepsilon]^T[\varepsilon]\right)}{\partial[\beta]} = \frac{\partial\left([R]^T[R]\right)}{\partial[\beta]} - \frac{\partial\left(2[\beta]^T[X]^T[R]\right)}{\partial[\beta]} + \frac{\partial\left([\beta]^T[X]^T[X][\beta]\right)}{\partial[\beta]}$$

$$\doteq 0 - 2[X]^T[R] + [X]^T[X][b] + [b]^T[X]^T[X]$$

$$= -2[X]^T[R] + 2[X]^T[X][b] = 0$$

set equal to zero because this is a minimization problem.

Solve for $[b]$

$$\Longrightarrow [b] = \left([X]^T[X]\right)^{-1}[X]^T[R]$$

5. Conclusion your best estimate for $[\beta] = \begin{bmatrix} A \\ B \\ D \\ C \end{bmatrix}$ in a least squares sense is

$$\boxed{[b] = \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \left([X]^T[X]\right)^{-1}[X]^T[R]}$$

Reference:

See

Draper, Norman and Smith, Harry. Applied Regression Analysis, 2nd Edition, Wiley Interscience, 1981.

See ~~pages~~ chapter 2