

Quadratic Multidimensional Regression

Eduardo M. de Moraes

July 6, 2020

Considering a structured dataset consisting of N sets $\{\mathbf{X}_k, f_k, \sigma_{y_k}\}$, where \mathbf{X}_k is the feature vector of dimension \mathbb{D} and f_k the target behaviour. Lets assume a multinomial model:

$$f_k = a + b^i x_{(i,k)} + c^{ij} x_{(i,k)} x_{(j,k)} \quad (1)$$

where $x_{(k)}$ with $k = 1, 2, \dots, \mathbb{D}$, represents the k -th feature of the vector \mathbf{X}_k . Here, we are using the Einstein sum notation. The constants a , b^i and c^{ij} are the parameters with values we want to find.

A general totally symmetric matrix M of rank R and dimension \mathcal{D} has a number of independent components given by

$$\text{number of components} = \binom{\mathbb{D} + R - 1}{R}.$$

where $\binom{\cdot}{\cdot}$ is the binomial coefficient. From (1), we can see clearly that c^{ij} is the components of a symmetric matrix. Because of this, we can write:

$$\text{number of components of } c^{ij} = \binom{\mathbb{D} + 2 - 1}{2}$$

The total number of independent coefficients of (1) is:

$$\text{number of parameters} = \binom{\mathbb{D} - 1}{0} + \binom{\mathbb{D}}{1} + \binom{\mathbb{D} + 1}{2}. \quad (2)$$

For a general regression problem, where the maximum order of the multinomial is \mathbb{O} , the number of parameter is:

$$n_P = \sum_{k=0}^{\mathbb{O}} \binom{\mathbb{D} + k - 1}{k}.$$

The expression of the square of (1) is:

$$\begin{aligned}
f_k^2 = & a^2 + 2ab^i x_{(i,k)} + (2ac^{ij} + b^i b^j) x_{(i,k)} x_{(j,k)} + \\
& + (2ad^{ijl} + 2b^i c^{jl}) x_{(i,k)} x_{(j,k)} x_{(l,k)} + \\
& + (2b^i d^{jlr} + c^{ij} c^{lr}) x_{(i,k)} x_{(j,k)} x_{(l,k)} x_{(r,k)} \\
& + (2c^{ij} d^{lrs}) x_{(i,k)} x_{(j,k)} x_{(l,k)} x_{(r,k)} x_{(s,k)} + \\
& + (d^{ijl} d^{rst}) x_{(r,k)} x_{(s,k)} x_{(t,k)} x_{(i,k)} x_{(j,k)} x_{(l,k)}
\end{aligned}$$

Assuming that the data are uncorrelated, the Chi Squared function can be written as:

$$\chi^2 = \sum_k \frac{1}{\sigma_k^2} [f_k - f(x_{(k)})]^2$$

or

$$\chi^2 = \sum_k \left(\frac{1}{\sigma_k^2} f_k^2 - \frac{2}{\sigma_k^2} f_k f + \frac{1}{\sigma_k^2} f^2 \right). \quad (3)$$

Defining the functions

$$A := \sum_k \frac{1}{\sigma_k^2} \quad B_i := \sum_k \frac{1}{\sigma_k^2} x_{(i,k)} \quad C_{ij} := \sum_k \frac{1}{\sigma_k^2} x_{(i,k)} x_{(j,k)} \quad ; \quad (4)$$

$$D_{ijl} := \sum_k \frac{1}{\sigma_k^2} x_{(i,k)} x_{(j,k)} x_{(l,k)} \quad E_{ijlr} = \sum_k \frac{1}{\sigma_k^2} x_{(i,k)} x_{(j,k)} x_{(l,k)} x_{(r,k)} \quad (5)$$

and

$$H := \sum_k \frac{f_k}{\sigma_k^2} \quad I_i := \sum_k \frac{f_k}{\sigma_k^2} x_{(i,k)} \quad J_{ij} := \sum_k \frac{f_k}{\sigma_k^2} x_{(i,k)} x_{(j,k)} \quad Z := \sum_k \frac{f_k^2}{\sigma_k^2}, \quad (6)$$

the χ^2 can be written as:

$$\begin{aligned}
\chi^2 = & Z - 2aH - 2b^i I_i - 2c^{ij} J_{ij} + \\
& + a^2 A + 2ab^i B_i + (2ac^{ij} + b^i b^j) C_{ij} + \\
& + (2b^i c^{jk}) D_{ijk} + (c^{ij} c^{kr}) E_{ijk r}
\end{aligned}$$

Minimizing with respect to a , b 's and c 's, we have the following system of equations:

$$aA + b^i B_i + c^{ij} C_{ij} + d^{ijk} D_{ijk} = H \quad (7)$$

$$aB_v + (b^j) C_{vj} + c^{jk} D_{vjk} + d^{ijk} E_{vijk} = I_v \quad (8)$$

$$aC_{vw} + b^i D_{ivw} + c^{ij} E_{ijvw} + d^{ijk} F_{vwijk} = J_{vw} \quad (9)$$

In matrix notation, this system can be rewritten as

$$\begin{bmatrix} A & B_1 & \cdots & B_{\mathbb{D}} & C_{11} & \cdots & C_{\mathbb{D}\mathbb{D}} \\ B_1 & C_{11} & \cdots & C_{1\mathbb{D}} & D_{111} & \cdots & D_{1\mathbb{D}\mathbb{D}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ B_{\mathbb{D}} & C_{\mathbb{D}1} & \cdots & C_{\mathbb{D}\mathbb{D}} & D_{\mathbb{D}11} & \cdots & D_{\mathbb{D}\mathbb{D}\mathbb{D}} \\ C_{11} & D_{111} & \cdots & D_{\mathbb{D}111} & E_{1111} & \cdots & E_{11\mathbb{D}\mathbb{D}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ C_{\mathbb{D}\mathbb{D}} & D_{1\mathbb{D}\mathbb{D}} & \cdots & D_{\mathbb{D}\mathbb{D}\mathbb{D}} & E_{11\mathbb{D}\mathbb{D}} & \cdots & E_{\mathbb{D}\mathbb{D}\mathbb{D}\mathbb{D}} \end{bmatrix} \begin{bmatrix} a \\ b^1 \\ \vdots \\ b^{\mathbb{D}} \\ c^1 \\ \vdots \\ c^{\mathbb{D}} \end{bmatrix} = \begin{bmatrix} A \\ G^1 \\ \vdots \\ G^{\mathbb{D}} \\ H^{11} \\ \vdots \\ H^{\mathbb{D}\mathbb{D}} \end{bmatrix}. \quad (10)$$

The solution for the system is:

$$\begin{bmatrix} a \\ b^1 \\ \vdots \\ b^{\mathbb{D}} \\ c^1 \\ \vdots \\ c^{\mathbb{D}} \end{bmatrix} = \begin{bmatrix} A & B_1 & \cdots & B_{\mathbb{D}} & C_{11} & \cdots & C_{\mathbb{D}\mathbb{D}} \\ B_1 & C_{11} & \cdots & C_{1\mathbb{D}} & D_{111} & \cdots & D_{1\mathbb{D}\mathbb{D}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ B_{\mathbb{D}} & C_{\mathbb{D}1} & \cdots & C_{\mathbb{D}\mathbb{D}} & D_{\mathbb{D}11} & \cdots & D_{\mathbb{D}\mathbb{D}\mathbb{D}} \\ C_{11} & D_{111} & \cdots & D_{\mathbb{D}111} & E_{1111} & \cdots & E_{11\mathbb{D}\mathbb{D}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ C_{\mathbb{D}\mathbb{D}} & D_{1\mathbb{D}\mathbb{D}} & \cdots & D_{\mathbb{D}\mathbb{D}\mathbb{D}} & E_{11\mathbb{D}\mathbb{D}} & \cdots & E_{\mathbb{D}\mathbb{D}\mathbb{D}\mathbb{D}} \end{bmatrix}^{-1} \begin{bmatrix} A \\ G^1 \\ \vdots \\ G^{\mathbb{D}} \\ H^{11} \\ \vdots \\ H^{\mathbb{D}\mathbb{D}} \end{bmatrix}. \quad (11)$$

The parametric correlation matrix is:

$$\mathbb{C} = \begin{bmatrix} A & B_1 & \cdots & B_{\mathbb{D}} & C_{11} & \cdots & C_{\mathbb{D}\mathbb{D}} \\ B_1 & C_{11} & \cdots & C_{1\mathbb{D}} & D_{111} & \cdots & D_{1\mathbb{D}\mathbb{D}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ B_{\mathbb{D}} & C_{\mathbb{D}1} & \cdots & C_{\mathbb{D}\mathbb{D}} & D_{\mathbb{D}11} & \cdots & D_{\mathbb{D}\mathbb{D}\mathbb{D}} \\ C_{11} & D_{111} & \cdots & D_{\mathbb{D}111} & E_{1111} & \cdots & E_{11\mathbb{D}\mathbb{D}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ C_{\mathbb{D}\mathbb{D}} & D_{1\mathbb{D}\mathbb{D}} & \cdots & D_{\mathbb{D}\mathbb{D}\mathbb{D}} & E_{11\mathbb{D}\mathbb{D}} & \cdots & E_{\mathbb{D}\mathbb{D}\mathbb{D}\mathbb{D}} \end{bmatrix}^{-1}. \quad (12)$$

The variance on parameters are:

$$\sigma_a^2 = \mathbb{C}_{1,1} \quad \sigma_{b^i}^2 = \mathbb{C}_{1+i,1+i}$$

Errors Propagation

$$f = a + b^i x_{(i)} + c^{ij} x_{(i)} x_{(j)}$$

$$\begin{aligned}
\sigma_f^2 = & \left(\frac{\partial f}{\partial a} \right)^2 \sigma_a^2 + \sum_i \left(\frac{\partial f}{\partial b^i} \right)^2 \sigma_{b^i}^2 + \sum_i \sum_{j>i} \left(\frac{\partial f}{\partial c^{ij}} \right)^2 \sigma_{c^{ij}}^2 + \\
& + 2 \frac{\partial f}{\partial a} \frac{\partial f}{\partial b^i} \sigma_{ab^i} + 2 \frac{\partial f}{\partial a} \sum_i \sum_{j>i} \left(\frac{\partial f}{\partial c^{ij}} \sigma_{ac^{ij}} \right) + \\
& + \sum_k \sum_i \sum_{j>i} \frac{\partial f}{\partial b^k} \frac{\partial f}{\partial c^{ij}} \sigma_{b^k c^{ij}}
\end{aligned}$$

$$\begin{aligned}
\sigma_f^2 = & \sigma_a^2 + \sum_i x_{(i)}^2 \sigma_{b^i}^2 + \sum_i \sum_{j>i} x_{(i)}^2 x_{(j)}^2 \sigma_{c^{ij}}^2 + \\
& + 2x_{(i)} \sigma_{ab^i} + 2 \sum_i \sum_{j>i} x_{(i)} x_{(j)} \sigma_{ac^{ij}} + \\
& + \sum_k \sum_i \sum_{j>i} x_{(k)} x_{(i)} x_{(j)} \sigma_{b^k c^{ij}}
\end{aligned}$$

$$\begin{aligned}
\sigma_f^2 = & \sigma_a^2 + \sum_i \left[x_{(i)}^2 \sigma_{b^i}^2 + \sum_{j>i} x_{(i)}^2 x_{(j)}^2 \sigma_{c^{ij}}^2 + 2x_{(i)} \sigma_{ab^i} + \right. \\
& \left. + 2 \sum_{j>i} x_{(i)} x_{(j)} \sigma_{ac^{ij}} + \sum_k \sum_{j>i} x_{(k)} x_{(i)} x_{(j)} \sigma_{b^k c^{ij}} \right]
\end{aligned}$$

$$\begin{aligned}
\sigma_f^2 = & \sigma_a^2 + \sum_i \left[x_{(i)}^2 \sigma_{b^i}^2 + 2x_{(i)} \sigma_{ab^i} + \sum_{j>i} x_{(i)}^2 x_{(j)}^2 \sigma_{c^{ij}}^2 + \right. \\
& \left. + 2 \sum_{j>i} x_{(i)} x_{(j)} \sigma_{ac^{ij}} + \sum_k \sum_{j>i} x_{(k)} x_{(i)} x_{(j)} \sigma_{b^k c^{ij}} \right]
\end{aligned}$$