Chi Squared Regression

Suppose we have a dataset $(\mathbf{x}_i, y_i)$ for $i = 1, \ldots, n$. We can collect the inputs $\mathbf{x}_i$ into an $n \times p$ matrix $\mathbf{X}$ with $\mathbf{x}_i^T$ as the rows

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix}$$

and the outputs y_i into a vector $\mathbf{y}^T = (y_1, y_2, \ldots, y_n)$. We pick the coefficients $\boldsymbol{\beta}$ that minimizes χ^2 defined by

$$\chi^2 = \frac{1}{n}\sum_{i=0}^{n-1} \frac{\left(y_i - \widetilde{y}_i\right)^{2s}}{\sigma_i^2}$$

where $\widetilde{y}_i = \sum_{j=0}^{p-1} X_{ij}\beta_j$. Taking the gradient of χ^2 with respect to the parameters β_k and setting it to zero:

$$\frac{\partial \chi^2}{\partial \beta_k} = \frac{1}{n}\sum_{i=0}^{n-1} \frac{1}{\sigma_i^2}\frac{\partial}{\partial \beta_k}\left(y_i - \sum_{j=0}^{p-1} X_{ij}\beta_j\right)^2$$

$$= \frac{1}{n}\sum_{i=0}^{n-1} \frac{1}{\sigma_i^2} 2\left(y_i - \sum_{j=0}^{p-1} X_{ij}\beta_j\right)(-X_{ik}) = \frac{1}{n}\sum_{i=0}^{n-1} \frac{1}{\sigma_i^2} 2\left(\sum_{j=0}^{p-1} X_{ij}\beta_j - y_i\right)X_{ik}$$

$$= \frac{1}{n}\sum_{i=0}^{n-1} \frac{1}{\sigma_i^2} 2X_{ik}\sum_{j=0}^{p-1} X_{ij}\beta_j - \frac{1}{n}\sum_{i=0}^{n-1} \frac{1}{\sigma_i^2} 2X_{ik}y_i = 0$$

$$\sum_{i=0}^{n-1} \frac{1}{\sigma_i^2} X_{ik}\sum_{j=0}^{p-1} X_{ij}\beta_j = \sum_{i=0}^{n-1} \frac{1}{\sigma_i^2} X_{ik}y_i$$

Define a matrix $\mathbf{A}$ by the elements $A_{ik} \equiv X_{ik} / \sigma_i$ and a vector $\mathbf{b}$ by the elements

$b_i = y_i \,/\, \sigma_i$. Then we can write

$$\sum_{i=0}^{n-1} A_{ik} \sum_{j=0}^{p-1} A\beta_j = \sum_{i=0}^{n-1} A_{ik} b_i$$

which has the corresponding matrix equation

$$\mathbf{A}^T \mathbf{A} \boldsymbol{\beta} = \mathbf{A}^{\mathbf{T}} \mathbf{b}$$

If $\mathbf{A}^T \mathbf{A}$ is invertible we get the single solution

$$\boldsymbol{\beta} = \left(\mathbf{A}^T \mathbf{A}\right)^{-1} \mathbf{A}^{\mathbf{T}} \mathbf{b}$$