A technical note on constraints and invariance in generalized linear models

Iain D Currie

Heriot-Watt University

To keep things simple we will deal with the standard linear model with normal errors; the results apply equally to the case of a generalized linear model.

We start with the definition of the *null space* of a matrix $\boldsymbol{M}$.

Definition: The *null space* of the matrix $\boldsymbol{M}$ is the set

$$\mathcal{N}(\boldsymbol{M}) = \{\boldsymbol{v} : \boldsymbol{M}\boldsymbol{v} = \boldsymbol{0}\}. \tag{1}$$

Lemma 1: $\mathcal{N}(\boldsymbol{M}'\boldsymbol{M}) = \mathcal{N}(\boldsymbol{M})$.

Proof: Let $\boldsymbol{v} \in \mathcal{N}(\boldsymbol{M}'\boldsymbol{M})$. Then

$$\boldsymbol{M}'\boldsymbol{M}\boldsymbol{v} = \boldsymbol{0} \Rightarrow \boldsymbol{v}'\boldsymbol{M}'\boldsymbol{M}\boldsymbol{v} = 0 \Rightarrow \boldsymbol{M}\boldsymbol{v} = \boldsymbol{0} \Rightarrow \boldsymbol{v} \in \mathcal{N}(\boldsymbol{M}). \tag{2}$$

Conversely, let $\boldsymbol{v} \in \mathcal{N}(\boldsymbol{M})$. Then

$$\boldsymbol{M}\boldsymbol{v} = \boldsymbol{0} \Rightarrow \boldsymbol{M}'\boldsymbol{M}\boldsymbol{v} = \boldsymbol{0} \Rightarrow \boldsymbol{v} \in \mathcal{N}(\boldsymbol{M}'\boldsymbol{M}). \tag{3}$$

Hence $\mathcal{N}(\boldsymbol{M}'\boldsymbol{M}) = \mathcal{N}(\boldsymbol{M})$. □

We consider the standard regression model with model matrix $\boldsymbol{X}$

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\epsilon}, \quad \boldsymbol{X} \text{ is } n \times p, n > p, \text{rank}(\boldsymbol{X}) = p - q, q \geq 1. \tag{4}$$

An estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ satisfies the *normal equations*

$$\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\theta}} = \boldsymbol{X}'\boldsymbol{y}. \tag{5}$$

We now prove the basic invariance lemma which says that the fitted values in a linear model are invariants, in the following sense.

Lemma 2 (Invariance): Let $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ be any two solutions of the normal equations. Then

$$\boldsymbol{X}\hat{\boldsymbol{\theta}}_1 = \boldsymbol{X}\hat{\boldsymbol{\theta}}_2. \tag{6}$$

Proof: Since $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ satisfy the normal equations (5)

$$\begin{aligned} & \boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_2) = \boldsymbol{X}'\boldsymbol{y} - \boldsymbol{X}'\boldsymbol{y} = \boldsymbol{0} \\ \Rightarrow \quad & \hat{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_2 \in \mathcal{N}(\boldsymbol{X}'\boldsymbol{X}) = \mathcal{N}(\boldsymbol{X}) \quad \text{by Lemma 1} \\ \Rightarrow \quad & \boldsymbol{X}\hat{\boldsymbol{\theta}}_1 = \boldsymbol{X}\hat{\boldsymbol{\theta}}_2 \end{aligned}$$

and we are done. □

We illustrate the use of this result by looking at the AP-model and the CBD model with cohort effects (M6).

Example 1 The age-period model.

The model matrix $\boldsymbol{X}$ for the AP-model is

$$\boldsymbol{X} = [\boldsymbol{X}_x : \boldsymbol{X}_y] = [\mathbf{1}_{n_y} \otimes \boldsymbol{I}_{n_x} : \boldsymbol{I}_{n_y} \otimes \mathbf{1}_{n_x}],\ n_x n_y \times (n_x + n_y), \tag{7}$$

with rank $n_x + n_y - 1$. Hence the rank of $\mathcal{N}(\boldsymbol{X})$ is one. Let

$$\boldsymbol{n} = \begin{pmatrix} \mathbf{1}_{n_x} \\ -\mathbf{1}_{n_y} \end{pmatrix} \tag{8}$$

then it is easy to check that $\boldsymbol{X}\boldsymbol{n} = \mathbf{0}$ and so $\boldsymbol{n}$ spans $\mathcal{N}(\boldsymbol{X})$. Let $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ be two solutions of the normal equations for the AP-model. Then $\hat{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_2 \in \mathcal{N}(\boldsymbol{X})$ and so

$$\begin{aligned} \hat{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_2 &= k \begin{pmatrix} \mathbf{1}_{n_x} \\ -\mathbf{1}_{n_y} \end{pmatrix} \text{ for some } k \\ \Rightarrow \quad \hat{\boldsymbol{\alpha}}_1 - \hat{\boldsymbol{\alpha}}_2 &= k\mathbf{1}_{n_x}, \quad \hat{\boldsymbol{\kappa}}_1 - \hat{\boldsymbol{\kappa}}_2 = -k\mathbf{1}_{n_y}. \end{aligned}$$

□

The relationships between other coefficients in other models can be found in the same way, i.e. find a basis for the null space of the model matrix.

Example 2 The CBD model with added cohort effects (M6).

The model matrix $\boldsymbol{X}$ for this model is

$$\boldsymbol{X} = [\boldsymbol{X}_x : \boldsymbol{X}_y : \boldsymbol{X}_c] = [\boldsymbol{I}_{n_y} \otimes \mathbf{1}_{n_x} : \boldsymbol{I}_{n_y} \otimes (\boldsymbol{x} - \bar{x}\mathbf{1}_{n_x}) : \boldsymbol{X}_c],\ n_x n_y \times (n_x + 3n_y - 1), \tag{9}$$

where $\boldsymbol{x} = (1, \ldots, n_x)'$ and $\boldsymbol{X}_c$ is that portion of the model matrix which corresponds to the cohort effects. The rank of $\boldsymbol{X}$ is $n_x + 3n_y - 3$ and so the rank of $\mathcal{N}(\boldsymbol{X})$ is two. Let

$$\boldsymbol{n}_1 = \begin{pmatrix} \mathbf{1}_{n_y} \\ \mathbf{0}_{n_y} \\ -\mathbf{1}_{n_c} \end{pmatrix} \quad \boldsymbol{n}_2 = \begin{pmatrix} (\bar{x} - n_x)\mathbf{1}_{n_y} - \boldsymbol{y} \\ \mathbf{1}_{n_y} \\ \boldsymbol{c} \end{pmatrix} \tag{10}$$

where $n_c = n_x + n_y - 1$, $\boldsymbol{y} = (1, \ldots, n_y)'$ and $\boldsymbol{c} = (1, \ldots, n_c)'$. It is straightforward to check that $\boldsymbol{X}\boldsymbol{n}_1 = \boldsymbol{X}\boldsymbol{n}_2 = \mathbf{0}$ and that $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are linearly independent. Hence

$$\boldsymbol{N} = \{\boldsymbol{n}_1, \boldsymbol{n}_2\} \tag{11}$$

is a basis for $\mathcal{N}(\boldsymbol{X})$. Let $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ be two solutions of the normal equations for the CBD model with cohort effects. Then $\hat{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_2 \in \mathcal{N}(\boldsymbol{X})$ and so

$$\hat{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_2 = A\boldsymbol{n}_1 + B\boldsymbol{n}_2 \tag{12}$$

for some A and B. Equating coefficients we arrive at equations (5) in the blog. The values of A and B for any two particular constraint systems can then be found by fitting the appropriate regression models.

You can find the null bases for the age-period-cohort and the APCI models in the talk referenced at the end of the blog.