Design matrix X with less than full column rank

Consider the model,

$$y_{ij} = \mu + \alpha_i + \tau_j + \epsilon_{ij}, i = 1, 2, \ldots, I; j = 1, 2, \ldots, J,$$

for the response from the ith treatment in the jth block, say. This can be put in the usual linear model form: $Y = X\beta + \epsilon$ as follows:

$$\begin{pmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1J} \\ y_{21} \\ y_{22} \\ \vdots \\ y_{2J} \\ \vdots \\ y_{I1} \\ y_{I2} \\ \vdots \\ y_{IJ} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & \ldots & 0 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & 1 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 1 \\ 1 & 0 & 1 & \ldots & 0 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 1 & \ldots & 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & 0 & 1 & \ldots & 0 & 0 & 0 & 0 & \ldots & 1 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & 0 & 0 & \ldots & 1 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 1 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & 0 & 0 & \ldots & 1 & 0 & 0 & 0 & \ldots & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_I \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_J \end{pmatrix} + \begin{pmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \vdots \\ \epsilon_{1J} \\ \epsilon_{21} \\ \epsilon_{22} \\ \vdots \\ \epsilon_{2J} \\ \vdots \\ \epsilon_{I1} \\ \epsilon_{I2} \\ \vdots \\ \epsilon_{IJ} \end{pmatrix}.$$

Here, X does not have full column rank. For instance, the first column is proportional to the sum of the rest. Thus $X'X$ is singular, so the previous discussion does not apply. β itself is not estimable, but what parametric functions of β are estimable?

Result. For any matrix A, the row space of A satisfies $\mathcal{M}_C(A') = \mathcal{M}_C(A'A)$.

Proof. $Ax = 0$ implies $A'Ax = 0$. Also, $A'Ax = 0$ implies $x'A'Ax = 0$, so $Ax = 0$. Therefore the null space of A and $A'A$ are the same. Consider the orthogonal space and note $\text{Rank}(A'A) = \text{Rank}(A) = \text{Rank}(A')$. Further, since $A'Aa = A'b$ where $b = Aa$, $\mathcal{M}_C(A'A) \subset \mathcal{M}_C(A')$. Since the ranks (or dimensions) are the same, the spaces must be the same.

Theorem. Let $Y = \theta + \epsilon$ where $\theta = X\beta$ and $X_{n\times p}$ has rank $r < p$. Then
(i) $\min_{\theta\in\mathcal{M}_C(X)} ||Y - \theta||^2$ is achieved (i.e., least squares is attained) when $\hat{\theta} = X\hat{\beta}$ where $\hat{\beta}$ is any solution of $X'X\beta = X'Y$;
(ii) $Y'Y - \hat{\beta}'X'Y$ is unique for all nonzero Y.

Proof. (i) $X'X\beta = X'Y$ always has some solution (for β) since $\mathcal{M}_C(X'X) = \mathcal{M}_C(X')$. However, the solution is not unique since $\text{Rank}(X'X) = r < p$.

Let $\hat{\beta}$ be any solution, and let $\hat{\theta} = X\hat{\beta}$. Then $X'(Y - \hat{\theta}) = 0$. However, given $Y \in \mathcal{R}^n$, the decomposition, $Y = \hat{\theta} \oplus (Y - \hat{\theta})$ where $Y - \hat{\theta}$ is orthogonal to $\mathcal{M}_C(X)$ is unique, and for such a $\hat{\theta}$, $||Y - \theta||^2$ is minimized. We know from previous discussion that $\min_{\theta \in \mathcal{M}_C(X)} ||Y - \theta||^2$ is achieved with $\hat{\theta} = PY$ which is unique.
(ii) Note that

$$Y'Y - \hat{\beta}'X'Y = Y'Y - \hat{\theta}'Y = (Y - \hat{\theta})'(Y - \hat{\theta}),$$

since $\hat{\theta}'(Y - \hat{\theta}) = 0$. Also, $(Y - \hat{\theta})'(Y - \hat{\theta}) = ||Y - \hat{\theta}||^2$ is the unique minimum.

Question. Earlier we could find $\hat{\beta}$ directly. How do we find $\hat{\theta}$ now?

Projection matrices

From the theory of orthogonal projections, given $X_{n\times p}$ (i.e., p many n-vectors), there exists $P_{n\times n}$ satisfying
(i) $Px = x$ for all $x \in \mathcal{M}_C(X)$, and
(ii) if $\xi \in \mathcal{M}_C^{\perp}(X)$, then $P\xi = 0$.
What are the properties of such a P?

1. P is unique: Suppose P_1 and P_2 satisfy (i) and (ii). Let $w \in \mathcal{R}^n$. Then $w = Xa + b$, $b \in \mathcal{M}_C^{\perp}(X)$. Then,

$$(P_1 - P_2)w = (P_1 - P_2)Xa + (P_1 - P_2)b = (Xa - Xa) + (P_1 b - P_2 b) = 0.$$

Since this is true for all $w \in \mathcal{R}^n$, we must have $P_1 - P_2 = 0$.

2. P is idempotent and symmetric:

$$P^2 x = P(Px) = Px = x \text{ for all } x \in \mathcal{M}_C(X);$$

$$P^2 \xi = P(P\xi) = P0 = 0 \text{ for all } \xi \in \mathcal{M}_C^{\perp}(X).$$

Therefore P^2 satisfies (i) and (ii), and since P is unique, $P^2 = P$. Further, $Py \perp (I - P)x$ for all x, y, so that $y'P'(I - P)x = 0$. i.e., $P' = P'P$, so $P = (P')' = (P'P)' = P'P = P'$.

Result. Let Ω be a subspace of the vector space $\mathcal{R}^n$, and let P_Ω be its projection matrix. Then $\mathcal{M}_C(P_\Omega) = \Omega$.

Proof. Note that $\mathcal{M}_C(P_\Omega) \subset \Omega$. For this, take $y \in \mathcal{M}_C(P_\Omega)$. Then y is a linear combination of columns of P_Ω, or $y = P_\Omega u$ for some u. Since $u = w \oplus v$, $w \in \Omega$, $v \in \Omega^{\perp}$, we have, $y = P_\Omega u = P_\Omega(w \oplus v) = P_\Omega w = w \in \Omega$. Conversely, if $x \in \Omega$, then $x = P_\Omega x \in \mathcal{M}_C(P_\Omega)$.

$I_n - P_\Omega$ represents the orthogonal projection. i.e., $\mathcal{R}^n = \Omega \oplus \Omega^\perp$. Thus for any $y \in \mathcal{R}^n$, we have $y = P_\Omega y \oplus (I - P_\Omega)y$.

If $P_{n\times n}$ is any symmetric idempotent matrix, it represents a projection onto $\mathcal{M}_C(P)$: if $y \in \mathcal{R}^n$, then $y = Py + (I - P)y = u + v$. Note

$$u'v = (Py)'(I - P)y = y'P(I - P)y = y'(P - P^2)y = 0,$$

so that we get $y = u \oplus v$, $u \in \mathcal{M}_C(P)$, $v \in \mathcal{M}_C^\perp(P)$.