Orthogonal projections in multiple regression

stats545.1 Theory of linear models

Gunnar Stefansson

September 18, 2018

Subspaces and degrees of freedom

Assume $rank(\mathbf{X}) = r$

We have $y - \mathbf{X}\hat{\beta} \perp \mathbf{X}\hat{\beta}$ so that $y - \mathbf{X}\hat{\beta} \in \mathbf{V}^{\perp} = \{v : v \perp sp(\mathbf{X})\}$ and $dim(\mathbf{V}^{\perp}) = n - r$.

$$\underbrace{\hat{\mathbf{e}}}_{n\times 1} = y - \mathbf{X}\hat{\beta} = y - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$$

$$= (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')y = (\mathbf{I} - \mathbf{H})y$$

and $rank(\mathbf{I} - \mathbf{H}) = dim(\mathbf{V}^{\perp}) = n - p$

Assume $rank(\mathbf{X}) = r$ ($\mathbf{X}$ is $n \times p$ and usually $p = r$ but $r < p$ is common also).

The model $y = \mathbf{X}\beta + e_1$ is estimated using the projection $\hat{\beta} = (\mathbf{X}^{'}\mathbf{X})^{-1}\mathbf{X}^{'}y$ onto the subspace $sp(\mathbf{X})$.

We have $y - \mathbf{X}\hat{\beta} \perp \mathbf{X}\hat{\beta}$ svo $y - \mathbf{X}\hat{\beta} \in \mathbf{V}^{\perp} = \{v : v \perp sp(\mathbf{X})\}$ and $dim(\mathbf{V}^{\perp}) = n - r$.

$$\underbrace{\hat{\mathbf{e}}}_{n\times 1} = y - \mathbf{X}\hat{\beta} = y - \mathbf{X}(\mathbf{X}^{'}\mathbf{X})^{-1}\mathbf{X}^{'}y$$

$$= (\mathbf{I} - \mathbf{X}(\mathbf{X}^{'}\mathbf{X})^{-1}\mathbf{X}^{'})y = (\mathbf{I} - \mathbf{H})y$$

The multivariate normal and related distributions

A basis for the span of X

Orthonormal basis, $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ for $\mathbf{R}^n$:

Using Gram-Schmidt, first generate $\mathbf{u}_1, \ldots, \mathbf{u}_r$ which span $sp\{\mathbf{X}\}$, with $rank\{\mathbf{X}\} = r$ and the rest, $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_n$ are chosen so that the entire set, $\mathbf{u}_1, \ldots, \mathbf{u}_n$ spans $\mathbf{R}^n$.

$$\begin{aligned} \mathbf{X}\hat{\beta} &= \hat{\zeta}_1 \mathbf{u}_1 + \ldots \hat{\zeta}_r \mathbf{u}_r \\ \mathbf{y} &= \hat{\zeta}_1 \mathbf{u}_1 + \ldots \hat{\zeta}_r \mathbf{u}_r + \hat{\zeta}_{r+1} \mathbf{u}_{r+1} + \ldots + \hat{\zeta}_n \mathbf{u}_n \end{aligned}$$

Q-R decomposition

$$\mathbf{Q} := \left[\mathbf{u}_1 \vdots \mathbf{u}_2 \vdots \ldots \vdots \mathbf{u}_n\right]$$

is the $\mathbf{Q}$ in the Q-R decomposition of $\mathbf{X} = \mathbf{QR}$.

If

$$\mathbf{z} = \left(\hat{\zeta}_1, \hat{\zeta}_2, \ldots, \hat{\zeta}_n\right)$$

then

$$\mathbf{z} = \mathbf{Q}'\mathbf{y}$$

and hence

$$E\left[\mathbf{z}\right] = \mathbf{Q}'\mathbf{X}\beta$$

$$V\left[\mathbf{z}\right] = \mathbf{Q}'\sigma^2\mathbf{I}\mathbf{Q} = \sigma^2\mathbf{I}$$

Variances of coefficients

For each i we obtain

$$V\left[\hat{\zeta}_i\right] = \sigma^2$$

Normality and independence of coeffients

Note that $\hat{\zeta}_i$ are linear combinations of the various y_j since $\hat{\zeta}_i = \mathbf{u}_i \cdot \mathbf{y}$.

When the y_i are independent Gaussian random variables, $\hat{\zeta}_i$ have zero covariance and are thus also independent.

Each $\hat{\zeta}_i$ is a coordinate in an orthonormal basis, $\hat{\zeta}_i = \mathbf{y} \cdot \mathbf{u}_i$. When Y_i are independent Gaussian random variables, $\hat{\zeta}_i$ also become independent. As a result, the sums of squares are related to $\sigma^2 \cdot \chi^2$-distributions in a natural way.

Expected values of coefficients

For $i = r+1, \ldots, n$ we obtain

$$E\left[\hat{\zeta}_i\right] = 0$$

Sums of squares and norms

$$SSE(F) = ||\mathbf{y} - \mathbf{X}\hat{\beta}||^2 \quad = \sum_{i=p+1}^{n} \hat{\zeta}_i^2$$

Each $\hat{\zeta}_i$ is a coordinate in an orthonormal basis, $\hat{\zeta}_i = \mathbf{y} \cdot \mathbf{u}_i$. When y_i are independent Gaussian random variables, $\hat{\zeta}_i$ also become independent. As a result, the sum of squares is related to a $\sigma^2 \cdot \chi^2$-distribution in a natural way.

Degrees of freedom

$SSE(F)$has $n - r$ degrees of freedom.

References Neter, J., Kutner, M. H., Nachtsheim, C. J. and Wasserman, W. 1996. Applied linear statistical models. McGraw-Hill, Boston. 1408pp.