A basis for V=span(X) stats545.2 545.2 The multivariate normal distribution and projections in the linear model

Gunnar Stefansson

September 5, 2022

Gunnar Stefansson

A basis for V = span(X)

September 5, 2022 1/9

Subspaces

In the model $y = X\beta + e$, assume rank(X) = r where X is $n \times p$ and $r \leq p$

Recall $X\hat{\beta}$ is a project so $y - X\hat{\beta} \perp X\hat{\beta}$ so that $y - X\hat{\beta} \in V^{\perp} = \{v : v \perp \}$ sp(X) and $dim(V^{\perp}) = n - r$. If r = p, then:

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

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

and $rank(I - H) = dim(V^{\perp}) = n - p$ Assume rank(X) = r (X is $n \times p$ and usually p = r but r < p is common also). The model $y = X\beta + e$ is estimated using the projection $\hat{\beta} = (X'X)^{-1}X'y$ onto the subspace sp(X). We have $y - X\hat{\beta} \perp X\hat{\beta}$ svo $y - X\hat{\beta} \in V^{\perp} = \{v : v \perp sp(X)\}$ and $dim(V^{\perp}) = n - r$. $\hat{\boldsymbol{\beta}} = \boldsymbol{v} - \boldsymbol{X} \hat{\boldsymbol{\beta}} = \boldsymbol{v} - \boldsymbol{X} (\boldsymbol{X}' \boldsymbol{X})^{-1} \boldsymbol{X}'^{\Box \, \flat \ \boldsymbol{\delta}} \stackrel{\boldsymbol{\delta}}{=} \boldsymbol{v} = \boldsymbol{X} \boldsymbol{X} \boldsymbol{X} \boldsymbol{X}^{\top} \boldsymbol{X}^{\Box \, \flat \ \boldsymbol{\delta}} \stackrel{\boldsymbol{\delta}}{=} \boldsymbol{X} \boldsymbol{X}^{\top} \boldsymbol{X}^{\Box \, \flat \ \boldsymbol{\delta}} \boldsymbol{X}^{\top} \boldsymbol{X}^{\top}$

Gunnar Stefansson

September 5, 2022 2/9

A basis for the span of ${\sf X}$

Orthonormal basis, $\{u_1, \ldots, u_n\}$ for \mathbb{R}^n :

Using Gram-Schmidt, first generate u_1, \ldots, u_r which span $sp\{X\}$, with $rank\{X\} = r$ and the rest, u_{r+1}, \ldots, u_n are chosen so that the entire set, u_1, \ldots, u_n spans \mathbb{R}^n .

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

Q-R decomposition

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

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

lf

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

then

$$z = Q'y$$

and hence

$$E[z] = Q'X\beta$$
$$V[z] = Q'\sigma^2 |Q = \sigma^2|$$

3

・ 同 ト ・ ヨ ト ・ ヨ ト

Variances of coefficients

For each *i* we obtain

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

→

Image: A match a ma

э

Expected values of coefficients

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

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

Gunnar Stefansson

A basis for V=span(X)

September 5, 2022 6 / 9

3

イロト イヨト イヨト

Sums of squares and norms

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

Each $\hat{\zeta}_i$ is a coordinate in an orthonormal basis, $\hat{\zeta}_i = y \cdot u_i$. 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.

(日)

Normality and independence of coeffients

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

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, $\hat{\zeta}_i = y \cdot u_i$. 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.

Degrees of freedom

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

3

(日)

Copyright 2022, Gunnar Stefansson This work is licensed under the Creative Commons Attribution-ShareAlike License. To view a copy of this license, visit http://creativecommons.org/licenses/by-sa/1.0/ or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.