Estimability

Søren Højsgaard

October 1, 2013

Consider an n dimensional random vector y for which $\mathbb{E}(y) = \mu = X\beta$ ans $\mathbb{C}\mathrm{ov}(y) = V$. Here X is an $n \times p$ with $n > p$ matrix which does not necessarily have full rank. A least squares estimate for β is

$$\hat{\beta} = (X^\top V^{-1} X)^- X^\top V^{-1} y$$

where A^- is a generalized inverse of A. We are interested in contrasts of the form $c = \lambda^\top \beta$. An estimate of such a contrast is $\hat{c} = \lambda^\top \hat{\beta}$. Generalized inverses are not unique and therefore $\hat{\beta}$ is not uniquely estimated, and hence the contrasts are not uniquely estimated. On the other hand the fitted values $\hat{y} = X\hat{\beta}$ are uniquely indentified.

However, for some λs we always get the same estimated contrast. Such contrasts are said to be estimable. These contrasts can be described as follows: We can only learn about β through $X\beta$ so the only thing we can say something about is linear combinations $\rho^\top X\beta$. Hence we can only say something about $\lambda^\top\beta$ if there exists ρ such that $\lambda^\top\beta = \rho^\top X\beta$, i.e., if $\lambda = X^\top\rho$, that is, if λ is in the column space $C(X^\top)$ of $X^\top$. That is, if λ is perpendicular to all vectors in the null space $N(X)$ of X. To check this, we find a basis B for $N(X) \subset R^p$ which can be done using SVD:

Example:

```
> ff <- factor(rep(1:3, each=2))
> X  <- cbind(rep(1,6),model.matrix(~0+ff)); X
```

```
    ff1 ff2 ff3
1 1   1   0   0
2 1   1   0   0
3 1   0   1   0
4 1   0   1   0
5 1   0   0   1
6 1   0   0   1
```

```
> y  <- 1:6
```

```
> S <- svd(X)
> lapply(S, zapsmall)
```

```
$d
[1] 2.828427 1.414214 1.414214 0.000000

$u
           [,1]       [,2]       [,3]       [,4]
[1,] -0.4082483  0.5534366  0.1644422 -0.7071068
[2,] -0.4082483  0.5534366  0.1644422  0.7071068
[3,] -0.4082483 -0.1343072 -0.5615113  0.0000000
[4,] -0.4082483 -0.1343072 -0.5615113  0.0000000
[5,] -0.4082483 -0.4191294  0.3970691  0.0000000
[6,] -0.4082483 -0.4191294  0.3970691  0.0000000

$v
           [,1]       [,2]       [,3] [,4]
[1,] -0.8660254  0.0000000  0.0000000  0.5
[2,] -0.2886751  0.7826776  0.2325563 -0.5
[3,] -0.2886751 -0.1899391 -0.7940968 -0.5
[4,] -0.2886751 -0.5927385  0.5615405 -0.5
```

The basis of $N(X)$ is

```
> B <- S$v[, S$d/max(S$d) < 1e-10, drop=F]
> (B <- as.numeric(B/max(B)))

[1]  1 -1 -1 -1
```

Hence valid vectors λ must satisfy that $\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 = 0$.

```
> lam <- c(1, 1, 0, 0)
> sum( lam*B ) # Orthogonal

[1] 2.220446e-16

> (b.hat <- as.numeric(MASS::ginv(t(X)%*%X) %*% t(X) %*% y))

[1]  2.625 -1.125  0.875  2.875

> sum( lam * b.hat )

[1] 1.5
```

Take another basis for the mean space:

```
> X2 <- X
> X2[,3]<-0
> (b.hat2 <- as.numeric(MASS::ginv(t(X2)%*%X2) %*% t(X2) %*% y))

[1]  3.5 -2.0  0.0  2.0

> sum( lam * b.hat2 )

[1] 1.5
```

On the other other hand, the following does not produce a unique estimate:

```
> lam2 <- c(1, 0, 0, 0)
> sum( lam2 * b.hat )

[1] 2.625

> sum( lam2 * b.hat2 )

[1] 3.5
```