

## TWO-DIMENSIONAL PANEL SS DECOMPOSITION

There are a few typos and some false equalities in Nerlove's appendix, obscuring the point. (It looks as though he might have changed his mind about notation midway through, but neglected to make the changes consistently.) Here, I'll mimic his notation, in particular by not using bold notation for matrices and vectors. However, I will continue to use the "dot subscript" notation for means selectively, i.e. whenever it is helpful to distinguish group means from period means. As a running example, I will work out the case of  $N = 3$  and  $T = 2$ .

Using the same basic algebra as in the ANOVA sums of squares decompositions that I did for the first lecture, one can show that the total variability in some variable  $x$  observed in  $N$  units  $T$  times each can be decomposed three ways:

$$\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})^2 = \begin{cases} \sum_i \sum_t ((x_{it} - \bar{x}_{i\cdot})^2 + (\bar{x}_{i\cdot} - \bar{x})^2) \\ \sum_i \sum_t ((x_{it} - \bar{x}_{\cdot t})^2 + (\bar{x}_{\cdot t} - \bar{x})^2) \\ \sum_i \sum_t ((x_{it} - \bar{x}_{i\cdot} - \bar{x}_{\cdot t} + \bar{x})^2 + (\bar{x}_{i\cdot} - \bar{x})^2 + (\bar{x}_{\cdot t} - \bar{x})^2) \end{cases}$$

where (as indicated above),  $\bar{x}_{\cdot t} = \frac{1}{N} \sum_i x_{it}$  is the period mean for period  $t$ ,  $\bar{x}_{i\cdot} = \frac{1}{T} \sum_t x_{it}$  is the unit (or "group" or "individual" depending on context) mean for unit  $i$ , and  $\bar{x} = \frac{1}{NT} \sum_i \sum_t x_{it}$  is the grand mean.

We wish to write out these identities in matrix notation. The matrices involved will subsequently be useful in developing the error-components model (random effects).

1. The total sum of squares can be written:

$$\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})^2 = x' [I_{NT} - \bar{J}_{NT}] x = x' V x,$$

where  $V = I_{NT} - \bar{J}_{NT}$  and  $\text{rank}(V) = NT - 1$ .

ex.

$$\begin{aligned} x' [I_6 - \bar{J}_6] x &= [x_{11} \ x_{12} \ x_{21} \ x_{22} \ x_{31} \ x_{32}] \begin{bmatrix} \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\ &= [(x_{11} - \bar{x}) \ (x_{12} - \bar{x}) \ \dots \ (x_{32} - \bar{x})] \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\ &= x_{11}^2 - \bar{x}x_{11} + \dots + x_{32}^2 - \bar{x}x_{32} \\ &= \sum_i \sum_t x_{it}^2 - \bar{x} \sum_i \sum_t x_{it} \\ &= \sum_i \sum_t x_{it}^2 - NT\bar{x}^2 = \sum_i \sum_t (x_{it} - \bar{x})^2 \end{aligned}$$

2. The within-group SS around the group means can be written:

$$\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2 = x' [I_{NT} - (I_N \otimes \bar{J}_T)] x = x' W_N x$$

where  $W_N$  can be re-written:

$$W_N = [I_{NT} - (I_N \otimes \bar{J}_T)] = [I_N \otimes (I_T - \bar{J}_T)]$$

and  $\text{rank}(W_N) = (T - 1)N$ .

ex.

$$\begin{aligned}
x' [I_6 - (I_3 \otimes \bar{J}_2)] x &= [x_{11} \ x_{12} \ x_{21} \ x_{22} \ x_{31} \ x_{32}] \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\
&= [(x_{11} - \bar{x}_1) \ (x_{12} - \bar{x}_1) \ \dots \ (x_{32} - \bar{x}_3)] \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\
&= x_{11}^2 - \bar{x}_1 \cdot x_{11} + x_{12}^2 - \bar{x}_1 \cdot x_{12} + \dots + x_{32}^2 - \bar{x}_3 \cdot x_{32} \\
&= x_{11}^2 - \bar{x}_1 \cdot x_{11} - \bar{x}_1 \cdot x_{12} + x_{12}^2 + \dots \\
&= x_{11}^2 - \bar{x}_1 \cdot (x_{11} + x_{12}) + x_{12}^2 + \dots \\
&= x_{11}^2 - \bar{x}_1 \cdot (2\bar{x}_1) + x_{12}^2 + \dots \\
&= x_{11}^2 - 2\bar{x}_1 \cdot \bar{x}_1 + x_{12}^2 + \dots \\
&= x_{11}^2 - 4\bar{x}_1 \cdot \bar{x}_1 + 2\bar{x}_1 \cdot \bar{x}_1 + x_{12}^2 + \dots \\
&= x_{11}^2 - 2\bar{x}_1 \cdot (x_{11} + x_{12}) + 2\bar{x}_1 \cdot \bar{x}_1 + x_{12}^2 + \dots \\
&= x_{11}^2 - 2\bar{x}_1 \cdot x_{11} + \bar{x}_1 \cdot \bar{x}_1 + x_{12}^2 - 2\bar{x}_1 \cdot x_{12} + \bar{x}_1 \cdot \bar{x}_1 \dots \\
&= \sum_i \sum_t (x_{it} - \bar{x}_i)^2
\end{aligned}$$

3. The between-unit SS (of group means around the grand mean) can be written:

$$\sum_{i=1}^N \sum_{t=1}^T (\bar{x}_i - \bar{x})^2 = x' [(I_N \otimes \bar{J}_T) - \bar{J}_{NT}] x = x' B_N x$$

where  $B_N$  can be re-written:

$$B_N = [(I_N \otimes \bar{J}_T) - \bar{J}_{NT}] = [(I_N - \bar{J}_N) \otimes \bar{J}_T]$$

and  $\text{rank}(B_N) = N - 1$ .

ex.

$$\begin{aligned}
x' [(I_3 \otimes \bar{J}_2) - \bar{J}_6] x &= [x_{11} \ x_{12} \ x_{21} \ x_{22} \ x_{31} \ x_{32}] \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\
&= \left[ \left( \frac{x_{11} + x_{12}}{2} - \frac{\sum_i \sum_t x_{it}}{6} \right) \ \dots \ \left( \frac{x_{31} + x_{32}}{2} - \frac{\sum_i \sum_t x_{it}}{6} \right) \right] \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\
&= [(\bar{x}_{1.} - \bar{x}) \ (\bar{x}_{1.} - \bar{x}) \ \dots \ (\bar{x}_{3.} - \bar{x})] \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\
&= \sum_i \sum_t (\bar{x}_{i.} - \bar{x})^2
\end{aligned}$$

4. The within-period SS (around the period mean) can be written:

$$\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_{.t})^2 = x' [I_{NT} - (\bar{J}_N \otimes I_T)] x = x' W_T x$$

where  $W_T$  can be re-written:

$$W_T = [I_{NT} - (\bar{J}_N \otimes I_T)] = [(I_N - \bar{J}_N) \otimes I_T]$$

and  $rank(W_T) = T(N - 1)$ .

ex.

$$\begin{aligned}
x' [I_6 - (\bar{J}_3 \otimes I_2)] x &= [x_{11} \ x_{12} \ x_{21} \ x_{22} \ x_{31} \ x_{32}] \begin{bmatrix} \frac{2}{3} & 0 & -\frac{1}{3} & 0 & -\frac{1}{3} & 0 \\ 0 & \frac{2}{3} & 0 & -\frac{1}{3} & 0 & -\frac{1}{3} \\ -\frac{1}{3} & 0 & \frac{2}{3} & 0 & -\frac{1}{3} & 0 \\ 0 & -\frac{1}{3} & 0 & \frac{2}{3} & 0 & -\frac{1}{3} \\ -\frac{1}{3} & 0 & -\frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & -\frac{1}{3} & 0 & -\frac{1}{3} & 0 & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\
&= [(x_{11} - \bar{x}_{.1}) \ (x_{12} - \bar{x}_{.2}) \ \dots \ (x_{32} - \bar{x}_{.2})] \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\
&= \sum_i \sum_t (x_{it} - \bar{x}_{.t})^2
\end{aligned}$$

5. The between-period SS (of period means around the grand mean) can be written:

$$\sum_{i=1}^N \sum_{t=1}^T (\bar{x}_{.t} - \bar{x})^2 = x' [(\bar{J}_N \otimes I_T) - \bar{J}_{NT}] x = x' B_T x$$

where  $B_T$  can be re-written:

$$B_T = [(\bar{J}_N \otimes I_T) - \bar{J}_{NT}] = [\bar{J}_N \otimes (I_T - \bar{J}_T)]$$

and  $\text{rank}(B_T) = T - 1$ .

ex.

$$\begin{aligned}
x' [(\bar{J}_3 \otimes I_2) - \bar{J}_6] x &= [x_{11} \ x_{12} \ x_{21} \ x_{22} \ x_{31} \ x_{32}] \begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\
&= [(x_{.1} - \bar{x}) \ (x_{.2} - \bar{x}) \ \dots \ (x_{.2} - \bar{x})] \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\
&= \sum_i \sum_t (x_{.t} - \bar{x})^2
\end{aligned}$$

6. The residual SS (of within-period individual variability) can be written:

$$\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_{i.} - \bar{x}_{.t} - \bar{x})^2 = x' [I_{NT} - (I_N \otimes \bar{J}_T) - (\bar{J}_N \otimes I_T) + \bar{J}_{NT}] x = x' W^* x$$

where  $W^*$  can be re-written:

$$W^* = [I_{NT} - (I_N \otimes \bar{J}_T) - (\bar{J}_N \otimes I_T) + \bar{J}_{NT}] = [(I_N - \bar{J}_N) \otimes (I_T - \bar{J}_T)]$$

and  $\text{rank}(W^*) = (N-1)(T-1)$ .

ex.

$$\begin{aligned} x' [I_6 - (I_3 \otimes \bar{J}_2) - (\bar{J}_3 \otimes I_2) + \bar{J}_6] x &= \\ [x_{11} \ x_{12} \ x_{21} \ x_{22} \ x_{31} \ x_{32}] &\begin{bmatrix} \frac{2}{6} & -\frac{2}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \\ -\frac{2}{6} & \frac{2}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} & \frac{2}{6} & -\frac{2}{6} & -\frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{2}{6} & \frac{2}{6} & \frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & \frac{2}{6} & -\frac{2}{6} \\ \frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & -\frac{2}{6} & \frac{2}{6} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\ &= \left[ \left( x_{11} - \frac{x_{11} + x_{12}}{2} - \frac{x_{11} + x_{21} + x_{31}}{3} + \frac{x_{11} + \dots + x_{32}}{6} \right) \dots \right] \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\ &= [(x_{11} - \bar{x}_{1.} - \bar{x}_{.1} + \bar{x}) \dots (x_{32} - \bar{x}_{3.} - \bar{x}_{.2} + \bar{x})] \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{bmatrix} \\ &= \sum_i \sum_t (x_{it} - \bar{x}_{i.} - \bar{x}_{.t} - \bar{x})^2 \end{aligned}$$

Next, note that:

$$V = \begin{cases} W_N + B_N \\ W_T + B_T \\ W^* + B_N + B_T \end{cases}$$

and

$$W_N B_N = W_T B_T = W^* B_N = W^* B_T = B_N B_T = 0$$



they also obtain. It is also easy to show that  $V, W_N, B_N, W_T, B_T$ , and  $W^*$  are all symmetric and idempotent.

R : Copyright 2004, The R Foundation for Statistical Computing  
Version 1.9.0 (2004-04-12), ISBN 3-900051-00-3

R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for a HTML browser interface to help.  
Type 'q()' to quit R.

```
> grun <- read.csv("c:/documents/teaching/493Panel/grfldat.csv")
> y <- grun[,3]
> i <- grun[,1]
>
> n <-max(i)
> t <-length(y)/n
>
> IN<-matrix(0,n,n)
> diag(IN)<-1
>
> IT<-matrix(0,t,t)
> diag(IT)<-1
>
> INT<-kronecker(IN, IT)
>
> JbN<-matrix(1/n,n,n)
> JbT<-matrix(1/t,t,t)
> JbNT<-matrix(1/(n*t),n*t,n*t)
>
> CN<-IN-JbN
> CT<-IT-JbT
>
> V <- INT-JbNT
> WN <-kronecker(IN,CT)
> BN <-kronecker(CN,JbT)
> WT <-kronecker(CN,IT)
> BT <-kronecker(JbN,CT)
> Wstar <-kronecker(CN,CT)
>
> TSS <- t(y) %*%V%*%y
> WGSS <- t(y) %*%WN%*%y
> BGSS <- t(y) %*%BN%*%y
> WPSS <- t(y) %*%WT%*%y
> BPSS <- t(y) %*%BT%*%y
> RSS <- t(y) %*%Wstar%*%y
>
> TSS
      [,1]
[1,] 9359944
> WGSS
      [,1]
[1,] 2244352
> BGSS
      [,1]
[1,] 7115592
> WPSS
      [,1]
[1,] 8731241
> BPSS
      [,1]
[1,] 628703.4
> RSS
      [,1]
```

Document2

11/6/2005

```
[1,] 1615649
>
> TSS-WGSS-BGSS
      [,1]
[1,] -4.656613e-09
> TSS-WPSS-BPSS
      [,1]
[1,] -2.328306e-09
> TSS-BGSS-BPSS-RSS
      [,1]
[1,] -2.793968e-09
>
>
> # some products are not exactly zero b/c of rounding
> # I won't demonstrate them b/c the matrices are so large,
> # but typing, e.g., WN%*%BN, etc. yields 200 x 200 matrices
> # all of whose elements are (something) x 10^-17 or less
> # ditto for V-WN-BN, V-WT-BT, V-Wstar-BN-BT
>
```

```

use c:/documents/teaching/493panel/grunf.dta

summarize i
gen ybardotdot = r(mean)

bysort firm: egen ybaridot=mean(i)
bysort year: egen ybardott=mean(i)

egen TSS = sum((i-ybardotdot)^2)
egen WGSS = sum((i-ybaridot)^2)
egen BGSS = sum((ybaridot-ybardotdot)^2)
egen WPSS = sum((i-ybardott)^2)
egen BPSS = sum((ybardott-ybardotdot)^2)
egen RSS = sum((i-ybaridot - ybardott + ybardotdot)^2)

list TSS WGSS BGSS WPSS BPSS RSS if _n==1

=====

clear

. do c:/stata/sumssqrs.do

. use c:/documents/teaching/493panel/grunf.dta

. summarize i

```

Variable	Obs	Mean	Std. Dev.	Min	Max
i	200	145.9583	216.8753	.93	1486.7

```

. gen ybardotdot = r(mean)

. bysort firm: egen ybaridot=mean(i)
. bysort year: egen ybardott=mean(i)

. egen TSS = sum((i-ybardotdot)^2)
. egen WGSS = sum((i-ybaridot)^2)
. egen BGSS = sum((ybaridot-ybardotdot)^2)
. egen WPSS = sum((i-ybardott)^2)
. egen BPSS = sum((ybardott-ybardotdot)^2)
. egen RSS = sum((i-ybaridot - ybardott + ybardotdot)^2)

. list TSS WGSS BGSS WPSS BPSS RSS if _n==1

```

	TSS	WGSS	BGSS	WPSS	BPSS	RSS
1.	9359944	2244352	7115592	8731241	628703.4	1615649

```

.
.
end of do-file

```